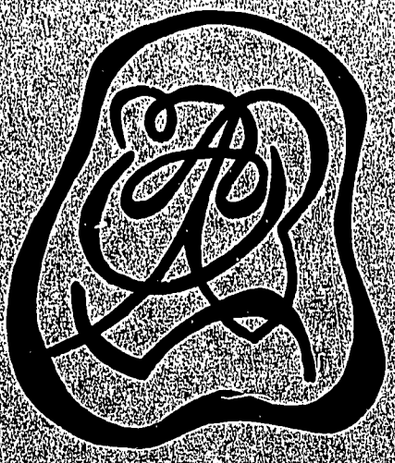
DOCUMENT RESUME

ED 053 216 TM 000 755

AUTHOR	Gibson, Dennis Lee; Dunnette, Marvin D.
TITLE	Predicting Individuals' Verbal Behavior in Counseling Group Interaction.
INSTITUTION	Minnesota Univ., Minneapolis.
SPONS AGENCY	Office of Naval Research, Washington, D.C. Personnel and Training Research Programs Office.
REPORT NO	TR-4006
PUB DATE	71
NOTE	71p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Behavior, Behavior Patterns, Behavior Rating Scales, College Students, *Comparative Analysis, Correlation, Group Behavior, *Group Dynamics, Homogeneous Grouping, Individual Activities, Interaction, *Interaction Process Analysis, Interpersonal Relationship, Measurement Techniques, Prediction, *Predictive Ability (Testing), Psychometrics, Self Evaluation, Testing, Tests, *Verbal Communication
IDENTIFIERS	*Hill Interaction Matrix, HIM

ABSTRACT

A self-report inventory was compared with a situational test as a predictor of the verbal behavior of individual members of small interpersonal skills training groups. As hypothesized, the situational test was a better predictor than was the self-report inventory. A powerful social conformity effect may have operated in both the situational test and the criterion groups, perhaps obscuring individual differences in preferred styles of interaction, which the self-report inventory seeks to measure. (Author/AG)



THE CENTER FOR THE STUDY OF ORGANIZATIONAL PERFORMANCE AND HUMAN EFFECTIVENESS

University of Minnesota
Minneapolis, Minnesota

Office of Naval Research Contract
ONR N00014-68-A-0141-0003

Approved for public release; distribution is unlimited.

PREDICTING INDIVIDUALS' VERBAL BEHAVIOR IN COUNSELING GROUP INTERACTION

Dennis Lee Gibson & Marvin D. Dunnette

Technical Report 4006

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
CENTER FOR THE STUDY OF ORGANIZATIONAL PERFORMANCE And HUMAN EFFECTIVENESS, University of Minnesota Minneapolis, Minnesota	Unclassified
	2b. GROUP

3. REPORT TITLE

PREDICTING INDIVIDUALS' VERBAL BEHAVIOR IN COUNSELING GROUP INTERACTION

4. DESCRIPTIVE NOTES (*Type of report and inclusive dates*)

Technical Report

5. AUTHOR(S) (*First name, middle initial, last name*)

Dennis Lee Gibson & Marvin D. Dunnette

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
1971	48	41

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N 00014-68-A-0141-0003	4006
b. PROJECT NO. NR 151-323	
c.	9b. OTHER REPORT NO(S) (*Any other numbers that may be assigned this report*)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
This is a revised version of Gibson's doctoral dissertation	Personnel & Training Research Programs Office of Naval Research (code 458) Department of the Navy Arlington, Virginia

13. ABSTRACT

A self-report inventory was compared with a situational test as a predictor of the verbal behavior of individual members of small interpersonal skills training groups.

The self-report predictor was a published instrument, the HIM-B (Hill Interaction Matrix). The situational test was a tape-recorded simulation of a small group meeting in which S imagined himself as a member and responded verbally at designated intervals. Subjects' responses were recorded and later rated on the Hill Interaction Matrix (HIM). Each predictor provided a profile of scores in the 16 cells of the HIM; these profiles were correlated with the criterion behavior profile for each of the Ss.

In total, 83 male college undergraduates were tested. Then 30 were selected to meet in six, five-man criterion groups, each of which met in two, 2-1/2 hour sessions with a trained leader. The Hill Interaction Matrix (HIM) was used to categorize the talk into a profile of criterion behavior for each member.

As hypothesized, the situational test was a better predictor than was the self-report inventory. Median correlations with Meeting #1 profiles were .72 and .03, respectively. Correlations with Meeting #2 were slightly lower. When cell scores on the self-report were differentially weighted as on the situational test, the median correlation with criterion profiles increased to .42. The weighted self-report was a

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

A-31408

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
	2b. GROUP
3. REPORT TITLE	
4. DESCRIPTIVE NOTES *(Type of report and, inclusive dates)*	
5. AUTHOR(S) *(First name, middle initial, last name)*	
6. REPORT DATE	7a. TOTAL NO. OF PAGES / 7b. NO. OF REFS
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	
10. DISTRIBUTION STATEMENT	
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

a significantly better predictor than was the unweighted, and the situational test was significantly better than either self-report.

Interpretations of these differences in predictive validity should be made with caution. A powerful social conformity effect may have operated in both the situational test and the criterion groups. This effect may have obscured individual differences in preferred styles of interaction, which the self-report inventory seeks to measure.

The six criterion groups were composed in different ways, based on HIM-B predictor profiles. Four groups were composed homogeneously, men who all scored high on some index on the HIM-B. Two groups were heterogeneous, in that the members in them had no high HIM-B score in common. Contrary to reports in the literature, prediction was no less accurate for members of heterogeneous groups than of homogeneous groups. The verbal behavior profiles of men in homogeneous groups changed more over time than did those of heterogeneous group members. Based on these correlational differences, and on clinical impressions, homogeneous groups were considered to be the more productive.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Group-to-Group Prediction						
Psychometric Prediction						
Group Composition						
Within-Group Behavior						
HIM - Hill Interaction Matrix						
HIM-B - Hill Interaction Matrix paper-and-pencil prediction instrument						
HIM-VG - Hill Interaction Matrix Vignette General, situational prediction instrument						
HIM-V4 - Hill Interaction Matrix Vignette, emphasis on Quadrant - unbalanced on weighted situational prediction instrument						
HIM-SS - Hill Interaction Matrix Statement-by-Statement interaction analysis system						
Early-to-Late group behavior correlations						
Small Group Verbal Behavior						
Self Report Inventory vs. Situational Test						
Interpersonal Interaction						

Predicting Individuals' Verbal Behavior in Counseling Group Interaction

Dennis Lee Gibson & Marvin D. Dunnette

April 1971

Center for the Study of Organizational Performance
& Human Effectiveness

University of Minnesota

Background

The purpose of this study is to compare some techniques devised for predicting the verbal behavior of individuals in small interpersonal skills training groups. In order to explore two fundamental areas in small group dynamics, viz., selection of members, and group composition, a method is needed to accurately predict the behavior of individual members of small groups.

The literature reviewed involves three general areas: prediction of individual behavior in a group; effects of composition on within-group behavior; and suggested methodology for prediction.

Prediction of Individual Behavior in a Group

An excellent review of literature done by Goldstein, Heller, and Sechrest (1966) is the backbone of the present literature review.

Group-to-Group prediction. The closest approximation to within-group behavior should logically be behavior in a preceding group. A number of studies in the group dynamics literature studied the consistency of individual behavior across a number of reconstituted groups.

Borgatta and Bales (1955a) studied the consistency of the observed behavior of 126 Air Force officers who met four times in groups of three men each. At each meeting the groups were reconstituted so that no S ever met twice with another S. A rater used Bales Interaction Process Analysis (Bales, 1950) to classify the observed behavior of each S in each meeting. The 126 Ss were studied in 14 groups of 9 men each.

The results of the study lead to the conclusion that 1) about 60% of the variance in a person's total volume of participation late in a meeting tends to arise from the same causes as it does earlier in the same meeting; and 2) that only about 30% of the variance in a person's total volume of participation in one meeting derives from the same sources as it does in another meeting with different partners.

In another study, Borgatta and Bales (1955b) looked at the characteristic rate of verbal behavior of each of 126 Ss, each participating in four different groups. They found that the characteristic rate of each individual affected the rates exhibited by other members with whom he met in a group. Although no correlations were reported, the experimenters concluded that they could use diagnostic sessions to predict how any particular combinations of Ss would interact with each other in later reconstituted groups.

Blake, Mouton, and Fruchter (1954) similarly used reconstituted 3-man groups of strangers. These groups met twice, each time working on a different task. Each S was rated by himself, by the two other group members, and by an observer, on variables such as his contribution to the discussion, leadership, and apparent frustration. The rankings were fairly consistent, with contingency coefficients in the range of .60 to .70 for some individuals. The authors concluded that a person's behavior tends to be consistent from group to group, and that it can be reliably judged by himself and by other observers.

Another study following a similar reconstituted-groups design (Bell and French, 1955) involved five-man groups, meeting six different times over a period of six weeks. The variable measured was leadership behaviors initiated. The authors concluded that individual characteristics account for about half of the variance in leadership status within the average group. They said, "Leadership status seems to be rather highly consistent despite the situational changes involved" (p. 279). In this context, a "situation" referred to a change in task and group composition.

In summary, the studies reviewed indicate that individual behavior is consistent enough from one group situation to another to allow predictions from one group to another. On a single index, such as rated leadership status, correlations from one session to another, across many persons, tend to run between about .50 and .90 (Bell and French, 1955; Borgatta and Bales, 1955a). The stability of such an index was said to be decreased by changes from one meeting to another in the composition and the assigned task of the groups.

Psychometric prediction to group behavior. Rather than using actual group behavior to predict later group behavior, a number of investigators have used various measures of personality to predict individual behavior in a group.

Toobert (1966) designed a research project to determine whether or not a personality measure could predict behavior equally well at two different points in time for the same individuals in the same groups. The task assigned to each group was discussion of a controversial subject. The personality measure used was the Guilford-Zimmerman Temperament Survey (GZTS). Bales' Interaction Process Analysis (IPA) was used to rate each S's group behavior. These IPA scores were rank-order correlated with the GZTS scores to give 140 correlations for the experimental group (the first meeting) and another 140 for the replication group (the second meeting of the same individuals). Only thirteen of the correlations were significant at the .05 level for the experimental group and nine were significant for the replication group. Toobert concluded that personality measures are not stable predictors of individual behavior in small groups. Individuals seemed to change with the nature of the group situation, regardless of their own personality characteristics.

Another study (Derr and Silver, 1962) used projective tests as predictors on 24 Ss engaged in group psychotherapy. These predictor scores were later correlated with the Ss' within-group behavior, as rated by their therapists. These correlations ranged from -.29 to +.43, averaged around .10, and were deemed insignificant. Derr and Silver saw their results as ". . . indirect evidence of the power of the group to shape and control the behavior of its members contrary to their personal predilections" (p. 324). They went on to suggest: "If such a finding is borne out by other research, it would point

to the futility of making predictions about an individual's behavior in a group, other than on the basis of what he does in a group similar or identical to a group in which his behavior is to be predicted" (pp. 324-325).

In their comprehensive review of 2699 references in the small group field, McGrath and Altman (1966) concluded (p. 108) that personality characteristics, measured by standard psychometric devices like the MMPI, are at most only slightly related to the content of interaction in a group and to interpersonal relations in the group. The same low relationship was said to be true of task abilities, and of attitudes toward issues, concepts and ideologies, such as authoritarianism.

Other studies, however, have yielded different results and conclusions. A study comparing an individual's early behavior in a group with later behavior in the same group was done by Kelly (1965). His measures of personality were the GZTS Sociability scale and the Ma scale of the Minnesota Multiphasic Personality Inventory (MMPI). High and low scores on these two scales were identified from among members of on-going counseling groups. Typescripts from the first and 21st meetings were rated on the Bales' IPA categories. Correspondence was noted between high tested scores and high rates of verbal behavior in the Bales' categories. No statistical measure of the degree of correspondence was reported in the abstract.

Mann (1959) used 100 male fraternity pledges in five-man groups in two reconstituted meetings. One meeting was given a task orientation; the other meeting featured social-emotional interaction. The three performance scales were Task Ability, Likability, and Tension. Mann found significant multiple correlations between the personality and performance scales. The R coefficients ranged from .30 to .50. Up to 25 percent of the variance in group behavior was therefore attributable to measured personality characteristics. The best single personality predictor variable was Social Extroversion. Although Mann did not explicitly compare group-to-group prediction with psychometric prediction, he did report that there were no significant differences in performance ratings between the two group conditions. Presumably behavior in one meeting might account for considerably more than 25% of the variance of behavior observed in the second meeting.

In a study that tends to support Mann's (1959) finding on Social Extroversion, Breer (1960) measured ascendance-submission among 25 college students, who met together in pairs. Ratings of the interaction observed in the pairs correlated .46 with predicted scores. Although Breer was not reporting multiple R's as Mann (1959) did, he found nearly the same proportion of variance (about 25%) in observed behavior predictable from self-report measures of a self-assertive personality trait.

In summary, the literature on psychometric prediction of group behavior reports some modest correlations (.30 to .50) between personality measures and observed group behavior (Breer, 1960; Mann, 1959). On the other hand, some writers report little if any significant relationship (Bennis, et al., 1957; Carr and Silver, 1962; McGrath and Altman, 1966; Toobert, 1966). The safest conclusion seems to be that some specific personality measures, such as Social Extroversion (Mann, 1959) and Ascendance-Submission (Breer, 1960), correlate much better with rated behavior in groups than do most general personality scales.

Summary. Some studies have reported comparisons of early behavior with later behavior on the part of individuals in small groups. Other studies have reported comparisons of psychometric data with observed behavior of individuals in groups. Correlation coefficients reported in the group-to-group, behavioral prediction studies (Bell and French, 1955; Blake, et al., 1954; Goldstein, et al., 1966) have generally run higher than those in the psychometric prediction studies (Derr and Silver, 1962; Breer, 1960; Mann, 1959). No studies were found in which behavior and psychometric prediction techniques were directly compared.

Effects of Group Composition on Individual Behavior

Most references to homogeneous grouping mention variables like intelligence, sex, and diagnostic category. Anderson (1969) called for research on different member selection variables: "Specifically, composing groups on the basis of predicted compatibility relative to preferred style of interaction appears most promising" (p. 212). Some studies that have been done along the lines Anderson suggested have shown some definite differences between homogeneous and heterogeneous groups.

In one study, known as the Harvard Compatibility Experiment (Schutz, 1966, pp. 128-136), members were selected on the basis of the need for affection scale. Twelve groups were formed: four homogeneous groups high on affection need, four homogeneous groups low on affection need, and four groups heterogeneous on the variable. The results showed that all homogeneous groups were significantly more productive on a problem-solving task than were heterogeneous groups, but there were no significant differences between high and low homogeneous groups.

In the study that is probably of greatest relevance to the present one, homogeneous groups seemed to allow members to behave in their preferred, natural manner, which was not true in groups that were heterogeneously composed (Gross, 1959). The composition variable used in Gross's study was a measure of personal-interpersonal orientation. The measure used to predict this variable was a specially-designed scale from the Fundamental Interpersonal Relations Orientation (FIRO-B) (Schutz, 1960). The measure used for observed, within-group behavior was the Hill and Hill Interaction Matrix, a forerunner of the present Hill Interaction Matrix (Hill, 1965). The Hill Interaction Matrix (HIM), like its forerunner, is a system for categorizing statements that are made in a therapy group.

In a validity study, Gross found evidence for validity of the FIRO-B as a predictor of within-group behavior. Twelve patients were rank-ordered on the personal-interpersonal variable based on their test scores. Two groups then met, one composed of the six patients with the highest Personal scores, and the other composed of the six patients with highest Interpersonal test scores. The patients were then rank-ordered on the personal-interpersonal variable based on their behavior in the groups. The rank-order correlation between the two sets of ranks was .87, significant beyond the .001 level. These twelve Ss had met in homogeneous groups. Based on later evidence presented by Gross, the correlation between tested and observed behavior would probably have been much lower if they had met in heterogeneous groups.

In his main experimental study, Gross (1959) selected, from a total sample of 125 mental patients, 12 who were Personal and 12 who were highly Interpersonal on their FIRO-B profiles. These Ss met in two homogeneous groups of six Personal subjects each and two homogeneous groups of six Interpersonal subjects each. In a counter-balanced design, the same Ss also participated in four heterogeneous groups, each composed of three Personal and three Interpersonal Ss. Each group met once for one hour, led by a therapist playing a neutral, innocuous role. Interaction analysis ratings were obtained from typescripts of 30 minutes near the end of each group meeting. Individuals in homogeneous groups spoke in the manner corresponding to their tested propensity. The same persons, meeting in heterogeneous groups, did not speak as much in their preferred styles, but rather, all talked in a common ground corresponding to nontherapeutic, socializing conversation. These observed styles of behavior were further confirmed by stimulated recall interviews conducted individually three to four hours after each group meeting.

Stager (1966) in a study of decision-making groups used conceptual level as a composition variable. Conceptual level refers to a style of information processing, which Stager measured for each individual, using a paper-and-pencil instrument. Four-man groups were composed with different percentages of high conceptual level members. Each group then was assigned a decision-making task. Significant differences were found with respect to how the groups interacted, and how they sought and used information. An implication for other studies on small groups was that noticeable effects arose from controlling group composition on a variable salient to the task of the group.

In a methodological paper, Magnusson, Gerzen, and Nyman (1968) took exception to some of the conclusions reached by other investigators. They described a "situation" as a combination of a task plus composition of a small group. They reported that changing either one, or neither task nor composition, still allowed high correlations among ratings of observed behavior in two group situations. However, changing both task and composition yielded totally random correlations between the two situations. In each case, the variables correlated were ratings of observed within-group behavior, made by two independent teams of judges. The authors concluded that if an individual's behavior in a small group is to "be regarded as an expression of the individual's general activity level, situational and interactional factors are of great importance" (p. 317).

In a survey of T-group research, Stock (1964) included a section on group composition (pp. 401-406). Her generalized conclusions were: " . . . group composition (based on certain personality variables) is a potent factor which finds rather direct expression in the character of the group interaction" (p. 405). "Homogeneous groups seem to reinforce and permit expression of the individual tendencies of the members, at least initially" (p. 406).

In summary, three tentative conclusions are indicated concerning composition effects:

1. Some investigators have reported that subjects exhibit their tested behavior only in groups composed homogeneously of persons with similar tested behavior. Therefore, any research on prediction of group

behavior should probably include some groups composed homogeneously on the variables under study (Gradolph, 1958; Gross, 1959; Stock, 1964).

2. The effects of composition are particularly noticeable when groups are composed on the basis of scores on variables salient to the task of the groups (Cecil, 1968; Gradolph, 1958; Stager, 1966).

3. Group composition is one important aspect to be controlled in an experimental situation. The other aspect is the task assigned to the group. If both these variables are changed, their impact may obscure the consistency of individual behavior across group meetings (Magnusson, et al., 1968).

Suggested Methodology in Prediction

In summarizing the results of 35 different groups on which research was done, Argyris (1968, pp. 192-193) had this rather dismal evaluation of self-report measures in prediction:

> . . . it was found that the participants were unable to predict their interpersonal behavior accurately ... If these data continue to be replicated, then the researchers who are studying interpersonal relationships may have to include observational data of the subjects' actual behavior because the interview or questionnaire data could be highly (but unknowingly) distorted.

Goldstein, et al. (1966) offered strong arguments to support their contention that group behavior should be predicted by behavioral measurement rather than self-report. They suggested the following research hypothesis: "On a variety of interactive communicative, and compatibility criteria, prediction of subsequent within-group behavior will be more accurate when based on direct behavioral measurement than on interview or psychometric measurement" (Goldstein, et al., 1966, p. 329). These authors went on to stress three criteria (p. 333) to be applied to situational testing as a behavioral measurement technique: (a) consistency, (b) relation to task success or outcome, and (c) objective observation.

The principle of consistency means that the situational test should approximate the real-life situation as nearly as possible. In studies of group counseling or group therapy, prediction based on trial groups would have the highest consistency; simulations would be next most preferred. Generalized self-report questionnaires would have least consistency. The principle of relation to task success is a reiteration of the suggestion to use predictor variables which are relevant to the raison d'etre of the group. Prior to development of the situational test, a definition of performance on the test must be made in terms directly comparable with measures of performance in the real-life criterion situation.

On the third criterion, objective observation, Goldstein, et al. (1966, p. 35) suggested several systems for interaction analysis of the behavior observed in criterion groups. The main point was that correct adherence to the first two

criterion would insure a proper predictive instrument; similar pains need to be taken with criterion measures. Criterion behavior ought to be assessed as objectively as possible, minimizing sources of unreliability between raters.

The three criteria above were drawn heavily from the report by Weislogel and Schwarz (1955). Their discussion of situational testing largely considered the work of the Office of Strategic Services (OSS) in World War II. The goal of the OSS studies was successful prediction of job success. If "job success" can be interpreted loosely to encompass real-life performance in general, situational testing such as done by the OSS should be of particular relevance to studies on therapeutic group interactions.

Anastasi (1968) observed that OSS situational tasks frequently showed low predictive validities. The validities were low because prediction was often from something quite specific, like building a structure with the help of uncooperative stooges, to very general real-life criteria such as advancement in military rank years later. In what might be of particular relevance to small group research, Anastasi cited the impressive validity of Leaderless Group Discussions (LGD): "Validity studies suggest that LGD techniques are among the most effective applications of situational tests" (Anastasi, 1968, p. 524). According to Anastasi, the LDG and other situational tests "appear to be most effective when they approximate actual work samples of the criterion behavior they are designed to predict" (1968, pp. 524-525).

From these references on suggested methodology, it appears that prediction of a person's group behavior should be based on his behavior in either a trial group or a suitable simulation of a real group. Criterion measures of actual, in-group behavior should use an objective system for interaction analysis. The interaction analysis system should demonstrate substantial inter-rater reliability, and it should tally frequencies of observed behaviors in its categories, rather than infer characteristics within the individuals being observed (Goldstein, et al., 1966, p. 335).

A pilot study was conducted to test the proposition that a situational test will predict group behavior better than will self-report devices. Subjects were five college men, all of whom were freshmen enrolled in an introductory sociology course at Augsburg College, a small Lutheran liberal arts school in Minneapolis. The paper-and-pencil instrument was the HIM-B, a 64-item questionnaire described more fully in Chapter II. The situational instrument was a tape-recorded simulation of some vignettes representing typical behavior in small counseling groups. This simulation was entitled the HIM-VG. The meaning of this name is "Hill Interaction Matrix--Vignette, General". Like the HIM-B, the HIM-VG consisted of 64 stimulus items, four for each of the sixteen cells of the Hill Interaction Matrix (HIM). (For a fuller explanation of the HIM framework, see Chapter II.)

Each S listened to the HIM-VG individually. After each vignette, a tone sounded, and the S spoke as he would if he had just heard that verbal exchange in a group of which he was a member. His response to each item was recorded and later rated into one of the 16 cells of the Hill Interaction Matrix (HIM).

For each S, then, predicted levels of behavior for each of the 16 cells of the Hill Interaction Matrix (HIM) were obtained by two methods, one psychometric and the other situational. Then, based on actual group behavior, an individual's response to the different types of stimuli represented by the 16 HIM cells, were determined for each S. This criterion behavior was measured by HIM-SS (Hill Interaction Matrix-Statement-by-Statement) ratings of the verbal interaction, individual-by-individual, in a one-hour meeting of the five men, with an experienced group leader. The experimenter rated all five HIM-VG's and the group meeting.

Statistical analyses consisted of intercorrelations of the 16 HIM cell scores obtained by each individual on the HIM-B and on the HIM-VG, with his HIM-SS group profile. There were five correlation coefficients (one corresponding to each group member) for HIM-VG vs. HIM-SS profiles. Both Pearson product-moment (PPM) and Spearman rank-order (RHO) correlation coefficients were calculated for each person. These correlations were O-type in Cattell's (1952) classification. (See Chapter II for an explanation of why significance tests cannot meaningfully be run on O-type correlations.)

As shown in Table 1, neither the HIM-B nor the HIM-VG predicted within-group HIM-SS behavior accurately. The HIM-B accounted on the average for about 10% of the variance in HIM-SS scores, but predicted in the opposite direction. The HIM-VG accounted for only about 4% of the variance in observed HIM-SS scores, and discriminated less clearly among individuals than did the HIM-B. The correlations of HIM-B and HIM-VG scores with each other averaged around zero. All three samples of behavior, the HIM-B, HIM-VG, and HIM-SS group behavior, appeared to be samples from different, unrelated domains of behavior.

One glaring observation from this pilot study was the unbalanced representation of HIM cells in the one hour of group talk. Both the HIM-B and HIM-VG prediction situations had balanced frequencies of stimulus statements, four for each of the 16 HIM cells. This balanced frequency obviously did not correspond with real-life behavior. The principle of consistency indicates that the situational test should be unbalanced in about the same way that observed behavior in the group is expected to be.

Another observation in the pilot study was the constricted behavior in the one-hour meeting. The warm-up consumed a great deal of time and represented only one or two categories of the HIM (cells 1 and 10).

The leader did considerable sponsoring of more valuable therapeutic interaction (cell 14), but verged on pressuring the group too much. It was apparent that the group would have to meet for a longer time in order for a wider range of verbal behavior to be elicited.

One earlier meeting of the five men without the leader produced one hour of talk in only three cells of the HIM. The need for a leader skilled in eliciting the full range of HIM behaviors was apparent. Seligman and Sterne (1969) reached a similar conclusion in their study of therapist-led and leaderless therapy groups. The leaderless group discussion in this pilot study also showed the need for speakers to be identified individually, since their voices could not be reliably differentiated in the tape recording.

Table 1

Correlations of predicted behavior with observed behavior in a pilot study group

	HIM-B vs HIM-SS		HIM-VG vs HIM-SS		HIM-B vs HIM-VG	
Member	PPM	RHO	PPM	RHO	PPM	RHO
1	-.36	-.30	.18	.43	.32	.05
2	.88	.80	.36	.02	.25	.26
3	-.35	-.54	.08	.24	-.38	-.32
4	.08	.37	.03	.00	.05	-.06
5	-.64	-.72	.34	.22	-.04	.00
Median	-.35	-.30	.18	.22	.05	.00
Range	-.64 to .88	-.72 to .80	.03 to .36	.00 to .43	-.38 to .32	-.32 to .26

In this pilot study no attempt was made to select the five men; their group was therefore heterogeneous on all tested variables. As some experimenters have reported (Gradolph, 1958; Gross, 1959), the members' true inclinations may not have been able to show through in a heterogeneous group, in which case near-zero predictive validities would be expected.

Another small study similar to the above pilot study was conducted by Thorn (1970). In Thorn's study, the two predictive instruments used were both paper-and-pencil: the HIM-B and a forced-choice reorganization of the 16 most highly-weighted items of the HIM-B. Subjects were male and female undergraduates in various years at Macalaster College, a small liberal arts school in St. Paul, Minnesota. Ss met in four-person groups for one hour to work on an assigned interaction task. A portion of each meeting was videotaped for later HIM-SS rating.

Thorn's results were similar to those obtained in Gibson's pilot study. Correlations of the two instruments with observed group behavior were in the range of .10 to .20. Even fewer HIM cells were used in these group discussions than were used in the Gibson pilot study. Again, it was evident that to elicit the full range of verbal behaviors represented by the 16 HIM cells, a leader was needed who could skillfully sponsor and model such behaviors. Since the higher-order cells in the HIM are of an intense interpersonal, therapeutic nature, an appropriate task for criterion group meetings would be a focus that might properly be termed relationship therapy or interpersonal skills training.

Implications of the Literature Review

No studies were found in the literature directly comparing situational and psychometric prediction of individual behavior in a small group. The literature does, however, suggest some conclusions pertinent to such a study:

1. Prediction of individual behavior in a group should be more accurate when done by a situational test than by a paper-and-pencil instrument (Argyris, 1968; Goldstein, et al., 1966).

2. The characteristics of the situational test should approximate the real-life criterion situation as nearly as possible (Anastasi, 1968; Goldstein, et al., 1966; Weislogel and Schwarz, 1955).

3. When criterion performance consists of ratings of the observed within-group behaviors of individuals, these ratings should be made with an objective system of interaction analysis. The interaction analysis system employed should show high inter-rater reliability, and power to discriminate between the behavior patterns of different individuals (Bennis, et al., 1957; Goldstein, et al., 1966).

4. Prediction of individual behavior may be more accurate for individuals participating in homogeneous rather than heterogeneous groups. The composition variable should be one relevant to the task of the group, and should also be reflected in the situational test (Gradolph, 1958; Gross, 1959).

5. If the setting for a predictive study is a therapy or human relations training group, instructions for the task would be, "Use your immediate interaction with each other to foster self-understanding in each member". This task should be carefully specified in both predictor and criterion situations.

6. The task appropriate to a therapy or human relations training group is not likely to be approached by a leaderless group of strangers who have had no previous group experience. A skilled leader is necessary to foster a wide range of interactions appropriate to the task (author's pilot study; Seligman and Sterne, 1969; Thorn, 1970).

7. Consistency of a situational test should be enhanced by making the frequency of each stimulus category represented in the overall test proportional to the expected frequency of the corresponding stimulus categories in the criterion situation (author's pilot study).

8. Criterion groups of strangers need to meet for a longer time than one hour in order for warm-up interactions to phase into task-relevant interaction. This conclusion is based on studies of four- and five-member heterogeneous groups. For groups of other sizes and compositions, a different conclusion may hold (author's pilot study; Thorn, 1970).

Design and Procedure

The purpose of this study was to compare a self-report inventory and a situational test as predictors of verbal behavior of individuals in counseling groups. The situational test was a tape-recorded simulation of a small group meeting, in which the subject being tested was asked to imagine himself as a participant. A number of college men were tested, and some were selected to meet in groups which were led by a trained group counselor. Some groups were homogeneous, some heterogeneous. All meetings were tape-recorded and then subjected to interaction analysis to provide the criterion behavior for each S against which his predicted behavior was correlated. The entire original pool of men was retested after the group meetings were concluded.

Subjects

Subjects for this study were 83 male college undergraduates drawn from a pool of students in an introductory psychology course at the University of Minnesota. Most of these men were 19 or 20 years old, and enrolled as sophomores in the College of Liberal Arts. They were recruited by personal phone calls and also by a voluntary sign-up sheet. They were told that they would all receive credit in the psychology course for their participation in the testing phase of the study. They were also told that only 30 men would be selected at random to meet in groups. These 30 would be paid $5 each and receive personal benefit from a five-hour interpersonal skills group experience. All were requested to be agreeable to either assignment: group participation or exclusion from group participation.

The Verbal Interaction Framework

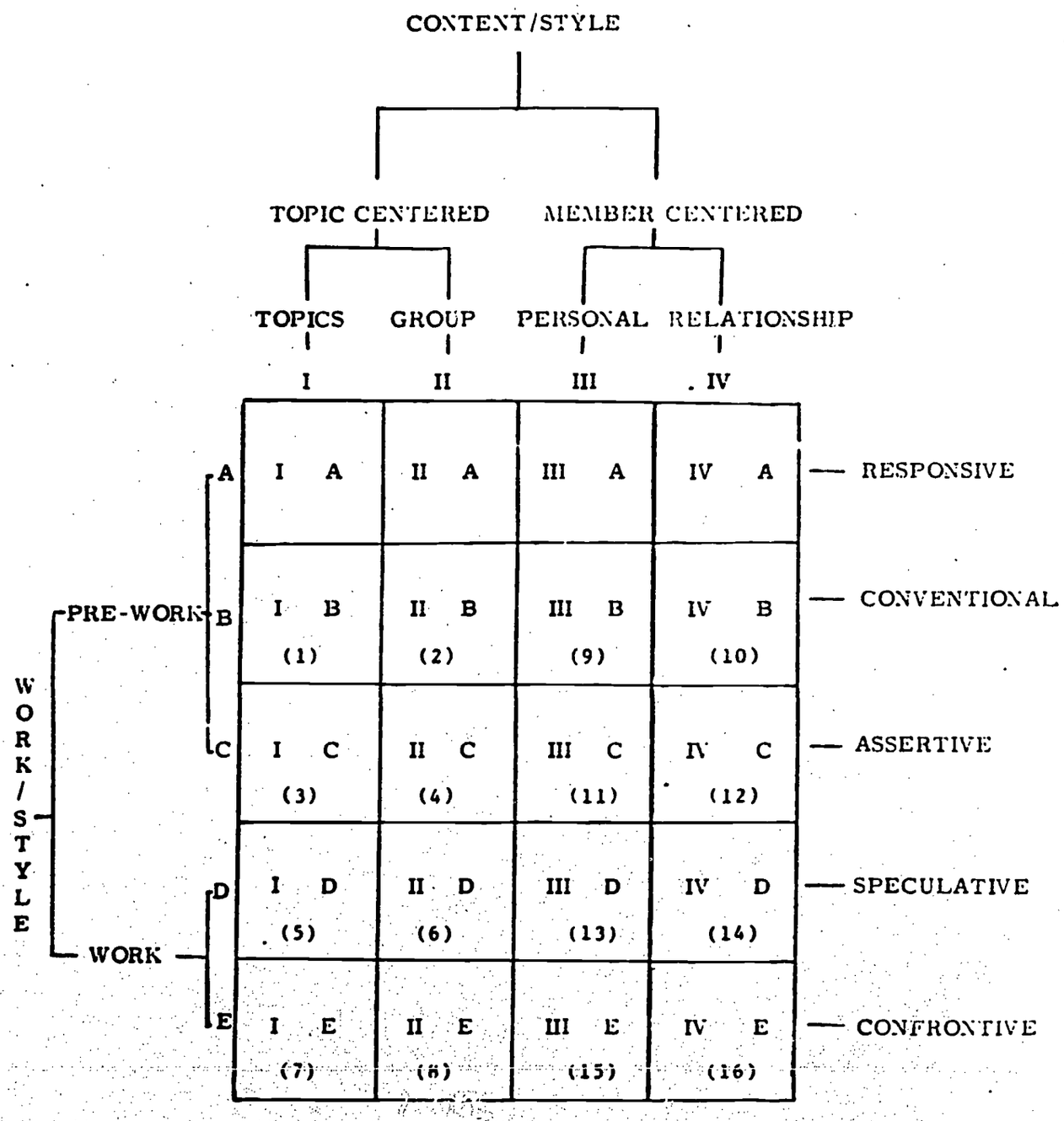
A system for verbal interaction analysis was required for categorizing the talk in both the behavioral predictor and in the later criterion groups. Any number of verbal interaction analysis systems could have been used. The Hill Interaction Matrix (HIM; Hill, 1965) was chosen because it has already spawned a paper-and-pencil psychometric instrument for predicting verbal behavior within the HIM framework. This instrument is the HIM-B (Hill, 1965, 1966).

The HIM categories are shown in Figure 1. The HIM categorizes a person's talk in two ways: first, what he talks about, and second, how he talks about it. The "What" dimension is shown by Roman numerals in Figure 1, where it is called "Content/Style". The safest thing a group can talk about is I, a topic of general interest, like the weather, politics, psychology, etc. Next, they can talk about the group itself (II). Next, they can participate in conversation that focuses on one present group member who is topic person. Such conversation is called Personal (III). The most risky thing to talk about, from the standpoint of vulnerability to embarrassment in the group, is IV, a relationship in the here-and-now, between two or more persons in the group.

So then, movement from left to right along the HIM content/style dimension is in the direction of greater interpersonal intimacy, and hence of greater therapeutic potency in Hill's theoretical value system (Hill, 1965).

HILL INTERACTION MATRIX

FIGURE 1



The way a person talks, the "How" dimension, is shown by A, B, C, D, and E in Figure 1, where it is called "Work/Style". Moving down on this dimension indicates an openness to changing one's opinions, attitudes, and characteristic behavior. Change such as this requires effort, so this "How" dimension is a "Work" scale. The A level, Responsive, refers to interactions that are very minimal and come only in response to great prodding by a therapist. Because A-level interaction is restricted largely to regressed hospitalized patients, it was not used in any of the interaction analysis in this study.

Conversation requiring the least effort in most groups is B, Conventional. This style of interaction is routine socializing, small talk, and where-are-you-from information-seeking. It takes only a little more effort to be Assertive (C). This is how a person talks when he argues, gripes, blows off steam, tells someone off, or tries to persuade. At neither B nor C is he willing to change anything about himself. He is not yet working, but is in "Pre-Work".

One begins to be open to change when he begins to think about the possibility of making a change. This thoughtfulness is reflected in D, the Speculative way of talking. The confrontive style, E, is the hardest work. It involves honesty, insight, taking responsibility for what is said by using specific examples, and getting down to the real core of the issue at hand.

The value system underlying the HIM is reflected in movement to the right and downward. This movement corresponds to increased interpersonal intimacy, and increased pressure toward change on the part of the participants. Hill (1965) has tentatively assigned a numeric rank from 1 to 16 to each of the cells. The higher the rank, the more therapeutically valuable the corresponding style of interaction should be. The upper left (IB) is cell 1, the lower right (IVE) is cell 16. Throughout the numerical data analysis in this study the cells were identified by their 1 to 16 rank.

When the interaction in a group is rated statement-by-statement on the Hill Interaction Matrix, the system used is called the HIM-SS. The HIM scoring manual (Hill, 1963) contains detailed rules for how to fit statements into one of the 16 cells. A number of illustrations are also presented with the rules for each cell. For the HIM-SS ratings done in this study, some supplementary scoring conventions were developed.

The HIM-SS has shown itself to be a reliable rating system in the hands of raters certified by the process described by Hill (1965, pp. 42-43). Inter-rater reliabilities were reported in three ways: (a) average percent agreement = 70%; (b) product-moment correlation = .76; and (c) rank-order correlation = .90 (Hill, 1965, p. 38).

Roffers (1969) used HIM-SS ratings of group interaction as a criterion measure of the effects of an operant conditioning treatment during group sessions. He said his results " . . . demonstrated that the Hill Interaction Matrix is a reliable and sensitive rating instrument by which relatively subtle aspects in alternative treatment procedures can be differentiated" (p. 141).

For the purposes of the present study, the HIM-SS seemed to have suitable reliability and objectivity as a measure of criterion within-group behavior.

Assessment Devices

The HIM-B. The HIM-B is a 64-item questionnaire designed to predict a person's preferred style of verbal interaction in a small group. A FORTRAN computer scoring program prints out a profile of scores such as appears in Appendix A. An interpretation to accompany the HIM-B profile is presented in Appendix B. This interpretation was prepared to hand to the experimental subjects, along with a copy of their printout. The HIM-B was used in this study both to select members for criterion groups, and to provide self-report prediction of individuals' within-group verbal behavior.

The HIM-V4. The behavioral predictive instrument that was compared with the HIM-B in this study was a situational test designated as the HIM-V4. The "V" stands for "vignette"; and "4" indicates the emphasis the instrument placed on interaction in Quadrant 4, the lower right-hand four cells of the HIM.

The HIM-V4 was a tape-recording of 70 verbal exchanges (i.e., vignettes) typical of what might occur in an interpersonal skills training group. A pool of vignettes was prepared partly from illustrations in the HIM scoring manual (Hill, 1963), partly from some illustrative items used to train HIM raters, and partly from the experimenter's own experiences in groups.

The man listening to the tape was asked to imagine himself as the fifth member of a group of which he was about to hear a simulated meeting. After each vignette, a tone was sounded and the subject was given 15 seconds in which to speak as he would if he were in the group and that verbal exchange had just taken place. His verbal response was then recorded on a second tape recorder to be rated later using the HIM-SS (Hill Interaction Matrix statement-by-statement interaction analysis system). Administration of the entire HIM-V4 took about 47 minutes.

The HIM-V4 in its development was preceded by an instrument used earlier in a pilot study, as described in Chapter I. This earlier version was called the HIM-VG, the "G" standing for "General". The HIM-VG was comparable to the HIM-B in that it consisted of 64 items, 4 for each of the 16 HIM cells. Experience in the pilot study led to development of the HIM-V4 as an unbalanced sampling of HIM behaviors, rather than a sampling of equal numbers of stimuli from each of the HIM cells, such as was true of the HIM-B and the HIM-VG. The six HIM cells chosen to be most heavily represented were the four in Quadrant 4, and cells III B and IV B. These six cells were emphasized because of their primary involvement in the type of interaction that was to be fostered in the interpersonal skills training groups. The particular distribution of the 70 stimulus items finally chosen for the HIM-V4 was set somewhat arbitrarily at 10 items for each of these six HIM cells, and one stimulus for each of the remaining 10 HIM cells. Some of the items were assigned to different cells than was the original intention, based on consensus among the three HIM raters.

The FIRO-B and Personality Research Form. The FIRO theory of interpersonal behavior (Schutz, 1960, 1966) is one of the most widely known conceptualizations of interpersonal behavior. The initials stand for Fundamental Interpersonal Relations Orientation. The FIRO-B is a 54-item questionnaire designed to predict an individual's orientation toward Schutz's three dimensions: Inclusion, Control, and Affection.

The Personality Research Form (PRF; Jackson, 1967) was selected as a general personality inventory whose scales might bear some relationship to style of interpersonal behavior.

Neither the FIRO-B nor the PRF was central to this study, and neither was involved in any analyses presented here. They were included in this study to provide potential data for future investigation of differential predictability and trainability of group members.

Stimulus-Response Analysis

In all published works using the Hill Interaction Matrix, the emphasis has been solely on categorizing verbal responses made by an individual or occuring within a group of unidentified individuals. Attention has not been paid to sequential analysis in the sense of identifying what stimuli elicit the rated responses. This stimulus-response analysis is characteristic of the Flanders system for interaction analysis, developed for use in classrooms (Amidon and Hough, 1967). A stimulus-response framework was considered appropriate in this study, especially to compare the situational test (the HIM-V4) with criterion group HIM-SS behavior.

As noted earlier in the literature review, prediction is expected to be most accurate for situational tests that follow the principle of consistency. That is, situations presented to the subject in the prediction task should be as similar as possible to the criterion situation. When prediction involves verbal interaction in a small group, it seems reasonable that the verbal responses compared in the predictor and criterion situations should be elicited by similar sets of stimuli. Similarity in this sense dictates proportional representation of the various stimulus categories in both situations.

In this study, verbal interaction was analyzed in terms of a 16 x 16, 256-cell stimulus-response matrix for each individual. The 16 response rows were the 16 HIM cells into which any given verbal response could be rated. A 17th response row was provided for unratable statements. This 17th row was not used in any predictive correlations. The 16 stimulus columns were again the 16 HIM cells. A stimulus was defined as the last ratable response made by the preceding speaker. The (I,J)th cell of this 256-cell matrix, therefore, represented the frequency of HIM type I statements the individual made in response to HIM type J stimulus statements.

In determining which statement was the stimulus for any given response, the following conventions were adopted:

1. A "speech" was defined as a duration of talk by one individual not interrupted by another speaker nor by a period of silence during which another person might have been expected to speak. A "statement" was defined as any portion of a speech to which a single HIM-SS rating was assigned. That is, HIM ratings were given to statements which could be as small as one word, or as large as one speech.

2. A speaker could not provide stimuli to himself. If he made five different statements during his speech, all five were taken as responses to the last ratable statement of the preceding speaker. His speech in effect created five new stimulus-response pairs.

3. Initial statements in a group were not scored as responses. They acted as stimuli to subsequent responses, but they themselves followed no definable stimulus. Similar non-scoring was accorded to statements that followed zero-level HIM ratings. Zero-level ratings were assigned to (a) statements understood by the responder but unintelligible to the HIM rater, and (b) long silences in which the prevailing agenda dissipated so as to make the next statement an initiating one.

Assessment Procedures

Prior to formation of the criterion groups, there were four testing sessions scheduled in an eight-day period. This testing period is referred to in this study as "pretesting". At each of the four pretesting sessions, the gathering of subjects was randomly split into two halves. One half was assigned to take the HIM-V4 first; the second half took the paper-and-pencil instruments first, the FIRO-B, the PRF, and the HIM-B. Then the HIM-V4 was presented to the second half. Those who did not complete their paper-and-pencil instruments before taking the HIM-V4 were allowed to do so afterwards. All four instruments were administered to each subject at one sitting.

After being pretested, 30 of the Ss were assigned to criterion groups. Meetings of these groups were held during a three-week period. Then all 83 Ss were retested. The reasons for retesting were two: (a) to get a measure of the test-retest reliabilities of the HIM-B and HIM-V4, and (b) to see if experience in a group caused subsequent tested behavior to be more realistic. That is, retesting the 30 Ss who met in groups would provide data for post-diction of group behavior. An instrument that gives high posttest correlations with behavior late in the life of a group may be useful in assessing the impact of the group experience on a member's characteristic behavior.

Pretest HIM-V4 recordings were defective for four group members: Member #4 of Group 1, Member #3 of Group 2, Member #4 of Group 2, and Member #4 of Group 3. The defects made these tapes impossible to rate. Postest HIM-V4 tape were rated for all five members of Group 4 and of Group 5, and also for five Ss who were not assigned to groups.

All of the HIM-V4 testing was administered by the experimenter. The paper-and-pencil tests were administered and scored by three hired female undergraduate clerks, except that the HIM-B's were computer-scored.

Subjects all received psychology course credit for the three hours they spent in testing. At the conclusion of the study, an interpretation of pretest HIM-B, FIRO-B, and PRF results was offered to any subject who requested it. About 10 or 12 men eventually asked for this interpretation.

Criterion Groups

Six groups of five men each were formed. Each group met for two, 2-1/2 hour sessions between the pretesting and retesting periods. The odd-numbered groups met on Tuesday evening and Saturday morning of the same week. The even-numbered groups met on Wednesday evening and Saturday afternoon. Two groups were run per week for each of the three consecutive weeks in February, 1970, the middle of the winter quarter. Each of the five members in each group was assigned a number 1 through 5. The leader was always assigned number 6. These numbers were used by an observer to identify the person speaking.

All interaction in each group was recorded on one track of a four-track stereo tape recorder. On the second track of the tape (the right channel), an observer simultaneously recorded the number of the person speaking. The observer was seated behind a one-way glass simply so that his presence and speaking would not be distracting to the group. At the beginning of each group, the leader explained fully to all members the purpose for the one-way glass, the observer's function, the goals of the experiment, and any other information requested by the members.

The leader was a male Ph.D. candidate in counseling psychology. He had extensive experience in both group and individual counseling in a variety of settings. He was also experienced as a practicum supervisor in both individual and group counseling training courses. He held a particularly strong theoretical orientation toward interpersonal skills training within the HIM framework. He was very conversant with the HIM, and with the goals and methodology of this study.

As mentioned previously, Gross (1959) studied the differences between groups composed homogeneously and groups composed heterogeneously on two measures of interpersonal orientation similar to the Personal (column III) and Relationship (column IV) Content Styles measured by the HIM-B. He found that participants in homogeneous groups exhibited their preferred style of interaction. In contrast, participants in heterogeneous groups did not exhibit their preferred style of interaction. Instead, they talked mostly in HIM cell IB, the only common ground they could find.

In the present study, a cue was taken from Gross's work, and an attempt was made to extend his study. Two of the six criterion groups were composed heterogeneously based on HIM-B profiles. The other four groups were composed of members homogeneously high on some index from the HIM-B. The HIM-B was used because it was easily scored and norms were available.

An index on the HIM-B refers to some linear combination of cell scores, such as a row, column or quadrant total, or a ratio of some of these totals. Beside each index printed by the HIM-B scoring program (see Appendix A) is a "+", "0", or "-" sign, designated as "norms." A "+" means that the value of the index for that person is higher than it was for 75% of the persons in the norm group. A "0" indicates the value is between the 25th and 75th percentile. A "-" corresponds to index values below the 25th percentile of the norm group.

The norm group used was 226 college undergraduates in Minnesota who took the HIM-B for various reasons during 1969 and 1970. Most of these persons were enrolled in sociology courses in which the HIM-B was used as one measure of the impact of the course experiences. The norm sample included both sexes. The 83 men in the present study were also included in the norm group.

In the terminology used here, "high" on a HIM-B index means that an individual had a "+" for that index on his HIM-B printout. "Average" corresponds to "0"; "low" refers to a "-".

Groups 1 and 5 were composed heterogeneously. That is, the five members in each did not share any high index score in common. Strictly speaking, Group 1 was homogeneously low on Row B (Conventional) and on Quadrant 1, and homogeneously average on Column III (Personal) and on Quadrant 3. Homogeneity on these variables, and at low and average levels was not seen as relevant to the intended Task of the criterion groups. The same comment applies to Group 5, which had all five members homogeneously low on Column I (Topic) and Row B (Conventional), and homogeneously average on Column III (Personal) and on Quadrant 4. On variables considered salient to interpersonal skills training, Groups 1 and 5 had no high preferences in common among the members composing each. According to Gradolph (1958) and Gross (1959), these groups would probably be characterized by strain and frustration.

Group 2 was homogeneous in a different sense from the other groups. It was composed of men whose highest quadrant score was in Quadrant 3, even though their Quadrant 3 scores may not have been high compared to the norm group. Actual interaction in Quadrant 3 is member-centered, and pre-work. The talk is about the persons present and it is either conspicuously friendly or conspicuously hostile, but does not seek to foster the self-understanding of the members.

One member of Group 2, member #5, was an exception of the Quadrant 3 composition rule. He was a last-minute replacement for a selected member who could not come. Member #5 had his highest quadrant score in Quadrant 4, and that score was "high". His score on Quadrant 3 was average. Since the goal of the leader was to foster Quadrant 4 interaction in the group, member #5 could be expected to function as a seed or catalyst to facilitate progress toward that goal. In summary, Group 2 was homogeneous on Quadrant 3 except for one member whose deviance was not expected to be disruptive to the group interaction.

Group 3 was homogeneously high on Risk Ratio, and on Row C (Assertive). Risk Ratio is explained in Appendix B. It is made high by a preference for Row C (Assertive) and Row E (Confrontive) ways of talking. In the case of Group 3, the uniform preference was for pre-work risk-taking of a hostile, provocative

kind. This group would be expected to show more hostility and assertion than any of the others.

Group 4 was homogeneously high on Work Ratio, which is explained in Appendix B. The members were also uniformly high on Row E (Confrontive) and on Quadrant 2. Quadrant 2 interaction is characterized by a cooperative effort to gain insight and understanding about topics relating to human behavior, and about the group itself. Of all six groups, Group 4 was selected so as to be already the most accomplished in the skills the group was intended to teach. Its interaction would be expected to work toward the gaining of insight about human interaction.

Group 6 was composed homogeneously high on Column IV (Relationship). It was originally intended to be composed on Member Ratio, which is explained in Appendix B. Four members were high on Member Ratio, but the fifth, a last-minute replacement for a drop-out, was low on Member Ratio. Since Group 6 was a homogeneous high Relationship group, it was very much like the groups that Gross (1959) designated as "Interpersonal". The expected natural interaction for this group would feature the discussion, exploration, and acting out of relationships among the members.

The groups were composed and numbers 1 through 6 assigned to them by a clerk, without the experimenter knowing which number was assigned to each group type. The experimenter was kept blind to this knowledge because he acted as observer of each of the group sessions and attempted to deduce the composition of each group from observed interaction. The group leader was not aware that any groups had been composed homogeneously. He and all participants were told that all group members were selected strictly at random.

All members were present for Meeting #1 of each group. Two of the 30 Ss were absent from Meeting #2: Member #4 of Group 2 and Member #3 of Group 3.

HIM Raters

Three raters were employed in this study, both to get a measure of inter-rater reliability, and also to share the rating load. All three were certified by the process outlined by Hill (1965) which requires at least 90% agreement with the HIM-SS ratings assigned by a panel of expert judges to a standard set of 64 written vignettes. All three raters had at least 60 hours of rating experience after certification.

The numbers assigned to the raters, 1 through 3, represented the rank of their sophistication in group processes, person 3 being the most sophisticated. They each rated one different group and the HIM-V4's for the members of that group. Each of the three raters also rated the last two hours from the first meeting of one of the groups, and the five pretest HIM-V4's for the men in that group. These ratings allowed several checks on inter-rater reliability: (a) on the 70 stimulus vignettes in the HIM-V4, (b) on the HIM-V4 profiles for each of five men, (c) on the HIM-SS group behavior profiles for each of five men plus the leader, and (d) on the HIM-SS group behavior profiles of one group as a whole.

In both the group ratings on HIM-V4 ratings, multiple responses to each stimulus were permitted. In the HIM-V4, up to three responses were scored following each stimulus.

Statistical Procedures

Computer programs were written to score the HIM-V4 and the HIM-SS group interaction in the form of stimulus-response matrices and several indexes of interaction. Another program was written to standardize corresponding HIM-V4 and HIM-SS profiles for a given individual. Copies of the standardized HIM-V4 and HIM-SS profiles for one S are presented in Appendix C. By "standardized" is meant that the number of stimuli of each type (each of the 16 HIM cells) was made equal in both the predictor and criterion situations. For example, if an individual was presented with ten stimulus statements of the cell 15 type in the HIM-V4, but only four stimulus statements of the cell 15 type in one HIM-SS time block, the correlation between his standardized HIM-V4 and HIM-SS profiles was based on the contribution of only four stimuli in cell 15. Since the smaller number of stimuli of the cell 15 type occurred in the HIM-SS profile, four responses to cell 15 stimuli were randomly selected from the ten in his HIM-V4 profile.

This profile standardization procedure was an attempt to correct for the unequal representation of stimulus items in predictor and criterion situations. A pilot study had revealed that individuals tended to give responses congruent to stimulus statements. The pilot study also showed that a mode of stimulus interaction could predominate such that an individual's underlying propensity was not brought into play. Standardization therefore allowed response profiles to be compared for a stimulus set common to two different situations.

Early-group HIM-SS time blocks were correlated with later HIM-SS time blocks to provide the maximum possible behavioral predictions. In this instance, within-group behavior was predicted by actual observed within-group behavior. Presumably no simulation could be expected to correlate as highly with HIM-SS behavior.

The correlations calculated among HIM-B, HIM-V4, and HIM-SS profiles were based on the total frequencies of responses in each of the 16 HIM cells. Both Pearson product-moment (PPM) and Spearman rank-order (RHO) correlations were calculated. These correlations were run both on raw score profiles, and on standardized profiles. For each prediction technique, one correlation was obtained for each S. For the most part, the median PPM and RHO coefficients for a number of Ss were about the same, and so only the PPM results were reported.

These correlations were O-type by Cattell's (1952) terminology. Each correlation was based on 16 variables (the scores on the 16 HIM cells), measured at two different occasions or by two different methods, for one person. This clarification is offered here to point out a possible difficulty in making inferences based on tests of significance of O-type correlations.[1] In the usual

[1] This caution was offered by Dr. David J. Weiss, University of Minnesota, in a personal communication.

R-type correlation, inference is made from a sample of persons to a population of persons of which the sample is representative. Analogous inference in the case of an O-type correlation that is significantly different from zero would be from a sample of variables to a population of variables of which the sample is representative. A representative sample of variables from a population of variables is difficult to conceptualize. Therefore, tests of significance on O-correlations are not reported in this study. What can be meaningfully reported are degrees of relationships between two methods as indicated by O-type correlation coefficients, and proportions of variance in common between two methods of measurement, as indicated by squared O-type correlations.

A significance test that could be meaningfully interpreted in this study was a correlated means t-test on differences between predictor-vs.-criterion correlations for two different prediction techniques. The procedure presented by McNemar (1962, p. 102) was followed. The two prediction techniques compared for each S were the HIM-B as a self-report device, and the HIM-V4 as a situational test. The criterion with which each was correlated was HIM-SS group behavior.

A third prediction technique was also compared with each of the other two, using the correlated means t-test. This technique was the weighted HIM-B. The 16 cell scores in each S's regular HIM-B profile were weighted by the frequency of stimuli in the corresponding cell on the HIM-V4. The resulting weighted HIM-B profile would then also reflect unbalanced or unequal cell frequencies. The reasoning in studying weighted HIM-B profiles was this:

Persons tend to respond with statements in HIM cells close to those in which they were addressed. Therefore, Ss taking the HIM-V4, with its unbalanced distribution of HIM stimulus types, tend to have unbalanced response distributions. Since similar unbalanced distributions are found in live group interaction, HIM-V4 profiles may correlate more highly with HIM-SS group behavior profiles than the regular HIM-B correlates with the HIM-SS simply because of unbalancing, and not by virtue of being a situational test. The HIM-B is unnecessarily handicapped by having the same number of stimuli in each HIM cell. The weighted HIM-B profile compared with the HIM-V4 profile should provide a more appropriate comparison of self-report and situational tests as prediction techniques.

In all correlations with HIM-SS group behavior profiles, a missing data check was included, which eliminated any HIM cell which did not occur as either a stimulus or a response in the HIM-SS profile. Without this missing data check, some correlations would be based partially on response frequencies of zero in some cells, whereas the prevailing talk in the group did not really allow these HIM cell responses.

Results and Conclusions

Reliability

Inter-rater. The 70 stimulus vignettes in the HIM-V4 provided a standard set against which to compare the three raters. In establishing the consensus ratings, there were 66 items which all three raters were able to rate. Rater #3 was unable to hear four of the stimulus items. Of the 66, there was unanimous agreement on 49% of the items or 32 items. At least two raters agreed on 61 of the 66 items of 93% of the items. The final consensus ratings were based on agreement between two or more raters. One of the five on which there was no agreement was eliminated.

Partly because three no-consensus items were awarded his ratings, Rater #2 showed the highest agreement with consensus ratings. Rater #1 showed the next highest, and Rater #3 the lowest agreement with consensus ratings. When the raters were compared with each other, number 1 and 2 had the highest agreement, Raters #2 and #3 the next highest, and Rater #1 with Rater #3 the lowest agreement. These conclusions held true for three measures of agreement: percent agreement, product-moment correlation (PPM), and rank-order correlation (RHO). The median values for each of these measures were respectively 74%, .78, and .79. Comparable values reported by Hill (1965, p. 38) were 70%, .76, and .90. Therefore, the three raters in this study showed inter-rater reliabilities on the HIM-V4 stimulus items about comparable to other reported reliabilities with the HIM-SS, except that the rank-order measure was slightly lower.

Another check on inter-rater reliability was afforded by the ratings of five pretest HIM-V4's of the members of Group 6. Table 2 presents these results. The median PPM and RHO reliabilities were .70 and .81 respectively. On the PPM measure, Raters #1 and #2 had the highest agreement, and #1 and #3 the lowest. On the rank-order measure, Raters #2 and #3 were highest, and Raters #1 and #2 were lowest.

Similar inter-rater reliabilites were calculated on HIM-SS response profiles for the last two hours of Meeting #1 of Group 6, which was rated by all three raters. Table 3 presents these correlations. On both PPM and RHO, Raters #1 and #2 showed the highest agreement and Raters #1 and #3 the lowest. The median PPM agreement was .59; the median RHO was .58.

In summary, inter-rater reliabilities for Raters #1 and #2 compared favorably with reliabilities reported elsewhere (Gibson, 1970; Hill, 1965). Rater #2 paired with Rater #3 showed slightly lower reliability, and Rater #1 paired with Rater #3 showed considerably lower reliability.

Inter-rater discrepancies. All three raters rated the five pretest HIM-V4's of Group 6 plus the last two hours of Meeting #1 of Group 6. These overlapping ratings allowed further comparisons of inter-rater reliability in the HIM system. All possible combinations of one rater's HIM-V4 with another rater's HIM-SS profile were correlated. The total number of correlations thus possible was 30. The median of these 30 cross-rater HIM-V4 vs. HIM-SS

Table 2

Inter-rater reliabilities on

5 HIM-V4 response profiles

(N=16 cells)

Group Member	RATER PAIRS 1 vs 2 PPM	1 vs 2 RHO	1 vs 3 PPM	1 vs 3 RHO	2 vs 3 PPM	2 vs 3 RHO
61	.84	.75	.68	.81	.70	.91
62	.68	.81	.73	.81	.84	.89
63	.88	.91	.69	.81	.79	.81
64	.87	.81	.47	.63	.73	.70
65	.63	.78	.50	.73	.65	.86
Median	.84	.81	.68	.81	.73	.86

Overall Median PPM = .70
Overall Median RHO = .81

Table 3

Inter-rater reliabilities on

7 HIM-SS response profiles

(N=up to 16 cells)

Group Member	RATER PAIRS 1 vs 2 PPM	1 vs 2 RHO	1 vs 3 PPM	1 vs 3 RHO	2 vs 3 PPM	2 vs 3 RHO
61	.95	.85	.32	.23	.46	.10
62	.42	.88	.30	.64	.89	.81
63	.75	.76	.70	.09	.58	.59
64	*	*	*	*	*	*
65	.84	.76	.20	.21	.60	.57
Leader	.98	.76	.44	-.04	.55	.56
Total Group	.79	.64	.48	.44	.70	.43
Median	.81	.76	.38	.22	.59	.56

Overall Median PPM = .59
Overall Median RHO = .58

* correlation not reported because based on fewer than 20 responses rated.

correlations was .45. The range was -.67 to .91. This way of comparing raters was probably the most exaggerated way to show inter-rater discrepancies. Spearman's correction for attenuation was applied to find the corrected median correlation of HIM-V4 with HIM-SS. These correlations once again were all product-moment O-type correlations based on raw profiles.

Median correlations used in the correction for attenuation were as follows: .45 for HIM-V4 vs. HIM-SS, .70 for HIM-V4 inter-rater reliability, and .59 for HIM-SS inter-rater reliability. The calculation is as follows:

$$.45 / (\sqrt{.70} \times \sqrt{.59}) = .70$$

This corrected median correlation agrees closely with the .72 median shown in Table 6. This level of correlation therefore appears to be a clear consensus among raters as to HIM-V4 correlations with HIM-SS observed behavior. Correlation that high was surprising. Judging by the "feel" of the HIM-V4's while rating them, the responses did not seem to be extremely valid reflections of the subjects' real-life behavior. Since all three raters knew that the HIM-V4 was expected to correlate highly with HIM-SS behavior, there is a possibility that part of that .70 correlation represents experimenter bias effects. The extent of this biasing, of course, is not determinable. In a replication of this study, an advisable procedure to minimize correlations due to bias on the part of raters would be to separate their ratings of HIM-V4 and HIM-SS meetings by several weeks. It was possible, when rating both HIM-V4's and group meetings within the same week or two, for a rater to remember the more colorful members. The biasing effect is probably not very great since Rater #2, who rated groups 1, 2, 3, and 6, over a period of four to five weeks, rated all group meetings first, and then went back and rated HIM-V4's. His correlations were not appreciably different from those of Rater #1 and Rater #3, who did their ratings in a shorter block of time. Rater #1 rated Group 4, and Rater #3 rated Group 5. These comparisons can easily be seen in Table 6 (pretest HIM-V with HIM-SS).

Situational test stability. The HIM-V4 test-retest correlations were calculated on the basis of individual item responses and on the basis of scores in the 16 cells. The overall median PPM correlation for the individual item responses was .54. When the test-retest correlations were calculated on the basis of scores in the 16 cells, rather than individual item responses, the correlations were greatly increased. The overall median was .91. The reason for the increased correlations on the 16-cell basis was the increased range of scores possible. The observed range of cell scores was 0 to 49, although the theoretical ceiling was 210. This increased range permitted greater correlations than did the restricted range of 1 to 16 which held true on the item basis. The appropriate test-retest correlation for the HIM-V4 is taken as .91. Subjects did not respond statement-by-statement in the same way in test and retest situations. However, their tendency toward a style of talk over the entire HIM-V4 was very similar in both situations. This style of talk during a block of time is what the 16-cell profile reflects. It seems to be the appropriate measure to correlate with talk during a period of time in a group.

Self-report stability. The HIM-B responses obtained in pretest and posttest situations were correlated in two different ways: On the basis of

item-by-item responses and cell-by-cell scores. In both cases, control subjects and experimental subjects had exactly the same median PPM correlations; for both groups the correlations were .60 for item-by-item analysis and .56 for cell-by-cell analysis. Rank-order correlations on the 16 cells were .58 and .59 for control and experimental Ss respectively.

Higher reliabilities were obtained on the weighted HIM-B profiles. The weightings used were the same as the frequencies that corresponded to the particular HIM stimulus cell on the HIM-V4. The median test-retest PPM correlation on 16 cells was raised from .56 for the unweighted HIM-B, to .88 for the weighted HIM-B. Again, no differences were found between control and experimental subjects. The increased reliability obtained by weighting is explained by the same reasoning mentioned earlier in connection with the HIM-V4. Differential weighting increased the range of cell scores, so that higher correlations could be obtained than with the restricted range of scores in the unweighted HIM-B. The unweighted HIM-B item scores could range from 0 to 10. On the weighted HIM-B, the possible range of cell scores was 0 to 110, which corresponds to the maximum cell weight (11), times the range of 0 to 10 for the unweighted HIM-B scores.

Because its variance more closely approximated that of HIM-SS profiles, the weighted HIM-B was considered more appropriate than the unweighted HIM-B, for correlations of self-report profiles with HIM-SS behavior. The test-retest reliability of the appropriate self-report device then was .88.

Early-to-Late Group Behavior Correlations

According to the reasoning in the design of this study, an individual's behavior in an actual group should be a better predictor of his later behavior in that group than should any self-report or situational test which attempts to simulate a group environment in a standardized fashion. A complicating factor in this reasoning is that, in a therapeutic group, individual behavior change is a goal of the group. Therefore, late behavior in an effective group should not necessarily have a high correlation with early behavior.

In the groups conducted in this study, the median correlation between raw-profile HIM-SS behavior in the first 2-1/2 hours, and the last 2-1/2 hours was .50. (Median product-moment and rank-order correlations on data have tended to be nearly the same, consequently, only PPM correlations are reported in the remaining tables.)

This rather modest correlation of .50 might be interpreted three ways: (a) the behavior of individuals in groups is so unstable that no predictor can be expected to correlate highly with observed behavior; or (b) in five hours therapeutic impact on the group was such that the members had behavior repertoires at the outset which contributed only 25% to the variance of their expanded behavior repertoires at the end; or (c) most members conformed with the prevailing style of talk in the groups, which was different for the two meetings.

Alternative (c) can be checked by removing the effects of changes in the prevailing style of talk during the two sessions. The concept of standardizing profiles, as explained in Chapter II, should allow Meeting #1 behavior to be compared with Meeting #2 behavior on a common basis. That is, standardized profiles for an individual would represent his responses to a similar set of stimuli at two different points in time.

Standardizing each individual's Meeting #1 and Meeting #2 HIM-SS profiles raised the median correlation to .82, compared with .50 for raw profiles. Alternative (c) above appears to hold true. Interpretations may now be recast:

1. Individuals in this study tended to react to similar stimulus patterns in about the same way late in the groups as they did early in the groups. About 65% of the variance in individual behavior rated early and late in the groups was attributable to the same sources. What these sources were is not clear, but some likely ones are discussed later.

2. The verbal behavior of the individuals in this study was stable enough that a good prediction technique might be expected to correlate as high as about .80 with observed behavior. That is, .80 is probably a ceiling on predictive validity.

3. The therapeutic impact of the groups in this study was not great enough in five hours to account for more than about 17% of the variance in individuals' behavior late in the groups. This estimate comes from applying Spearman's correction for attenuation, as explained by Helmstadter (1964, p. 84) to the .82 correlation between Meeting #1 and Meeting #2 standardized profiles. If HIM-SS ratings for Meetings #1 and #2 each had a reliability of .90 (a typical HIM-SS rater's test-retest reliability), the correlation between Meeting #1 and Meeting #2 measured perfectly would be $.82 / (\sqrt{.90} \times \sqrt{.90}) = .91$. This correlation of .91 represents about 83% of variance in common between Meeting #1 and #2 for the average group member. Therefore, no more than about 17% (and probably less) of the variance in an individual's behavior late in a group was due to changes that the group treatment itself may have caused in his characteristic interpersonal behavior.

Another way of looking at early-group behavior as a predictor of later-group behavior was a trend analysis carried out on Group 6. For each of the five men in this group, the HIM-SS profile for the first 1-1/4 hours was correlated with his HIM-SS profile for each of the subsequent three 1-1/4 hour blocks. The median PPM correlations were .26, .40, and .47. If the group behavior had been changing gradually, these correlations would have been expected to start high and descend regularly. Instead, they increased slightly within the moderate range. This cursory analysis showed no real trend of predictability. The HIM-SS profiles used in this trend analysis were raw rather than standardized profiles, which could not be obtained for the small number of statements made by each S in 1-1/4 hours. As in the correlations of Meeting #1 and Meeting #2, this finer trend analysis probably produced low correlations because of changes in the prevailing style of talk in each group, rather than because of gross instability of individual behavior or of therapeutic impact.

This finding presents a caution for research on verbal behavior in therapeutic groups. The actual talk pattern emitted during any portion of time in a group may vary considerably from that at any other portion of time, even from another portion very close in time. A person's observed talk in a group may have very little in common with his preferred style of talk. The way he actually talks may be largely determined by the moment-by-moment style of talk prevailing in the group. A rigorous analysis of predicted and observed behaviors should focus on the responses S makes to sets of stimuli which are similar in frequency and kind in both the prediction and criterion situations.

A rigorous analysis of this kind was provided in the present study by the profile standardizing procedure. However, this rigor of analysis does not diminish the need for caution in interpreting the results of analysis. The observation that persons in this study tend to adapt their talk to the prevailing style of talk in the group may point to a powerful social psychological effect operating in the groups. This effect may be something like conformity, or like a social desirability response set. Either of these phenomena might tend to make all persons respond alike to similar sets of stimuli, despite individual differences in preferred styles of talking. So then, some likely sources of variance in an indivdual's pattern of talk in a group are his personal preferences, his ability to discern expectations of others in the group, and his tendency to conform to whatever pressure these perceived expectations exert on him.

The finding in this study, that about 65% of the variance in rated behavior in one meeting is common with that in another meeting, seems roughly comparable to figures of 60% and 56% reported by Borgatta and Bales (1955a), and Bell and French (1955) in a leadership context. Strictly speaking, however, these figures cannot be compared, because they are based on entirely different types of correlations. The ones in this study were of patterns of talk by one person on two occasions (O-type correlations). They indicate the extent to which an individual distributes his talk over a fixed set of categories the same way on two different occasions. The correlations used in the leadership studies were of single indexes of behavior by many persons on two occasions (T-type correlations). These correlations indicate the extent to which persons on two different occasions tended to rank in the same order on one index. So then, O-correlations express profile or pattern stability; T-correlations express index stability.

Self-Report Prediction of Within-Group Behavior

Pretest HIM-B cell scores were correlated with HIM-SS cell scores for the verbal behavior of each group member in Meeting #1. As shown in Table 4, the median PPM correlation was .03, and the overall range was -.66 to .57. One interpretation might be that the HIM-B has no predictive validity. Another view is that, if the HIM-B is a valid measure of an individual's preferred pattern of talking in a group, some persons do behave congruently with their measured preferences (e.g., the .57 correlation), while others behave in ways distinctly opposite to their preferences (e.g., the -.66 correlation); and on the average most persons do not behave according to their preferences.

Table 4

Correlations of pretest HIM-B

with HIM-SS raw profiles

Group	Meeting #1 Median	Meeting #1 Range	Meeting #2 Median	Meeting #2 Range
1	.29	-.50 to .44	-.28	-.37 to .13
2	.06	-.47 to .24	-.20 (N=4)	-.51 to -.14
3	-.26	-.66 to .55	-.16 (N=4)	-.73 to .36
4	-.08	-.48 to .26	-.12	-.20 to .43
5	.00	-.19 to .31	-.22	-.47 to .21
6	.27	-.20 to .57	.11	-.40 to .25
Overall	.03 (N=30)	-.66 to .57	-.19 (N=28)	-.73 to .43
Homogeneous Groups	.12 (N-20)	-.66 to .57	-.10 (N=18)	-.73 to .43
Heterogeneous Groups	.00 (N=10)	-.50 to .44	-.25 (N-10)	-.47 to .21

Before arguing for either of these interpretations, attention must be paid to the caution made earlier. The 2-1/2 hour sample of behavior represented by Meeting #1 may have been heavily affected by the style of group interaction that prevailed during that time. This effect may have obscured the effect of individual preferences. An attempt must be made to equalize the frequency and kind of stimuli to which S responded in the predictor (HIM-B) and criterion (Meeting #1) situations.

A step in this corrective direction was taken by correlating weighted HIM-B profiles with HIM-SS profiles. As mentioned earlier, the weighted HIM-B profile for an individual consisted of his 16 HIM-B cell scores, each multiplied by the frequency of that cell's presentation in the HIM-V4 situational test. This set of weights was chosen as an approximation to the unbalanced distribution of stimuli in the actual groups, and also to allow later comparisons between the HIM-V4 and a similarly-weighted HIM-B. Table 5 shows that the weighted HIM-B had a median correlation of .42 with Meeting #1 behavior; which was significantly higher than the .03 median for the unweighted HIM-B.

The implication of this increased correlation is that self-report inventories such as the HIM-B might be made more useful predictors of the patterns of individuals' behavior in groups simply by scoring the responses in a way tailored to the situation to which prediction is being made. More weight should be given to scores in categories which are expected to be heavily used in the groups. Specifically, each HIM-B cell score should be multiplied by the expected frequency of stimulus statements in that HIM cell in an actual group meeting.

Tables 4 and 5 also show correlations separately for members of homogeneous and heterogeneous groups. Neither the HIM-B nor the weighted HIM-B as predictor showed any substantially different correlation with observed behavior in either type of group. The reports by Gradolph (1958) and Gross (1959) were not borne out in this study; members of homogeneous groups did not behave in greater accordance with their measured preferences than did members of heterogeneous groups. Gradolph's groups were leaderless and Gross explicitly instructed the therapist to be innocuous in his groups. By contrast, the leader in the present study was very active. This failure to replicate the findings of Gradolph (1958) and Gross (1959) may also be further evidence that the preferences of individuals in this study were obscured by more powerful effects operating in the groups, one of which was the leader's modeling and urging of particular categories of talk more than of others.

The median correlation of the pretest HIM-B with Meeting #2 HIM-SS profiles (Table 4) was -.19. The similar correlation for the weighted HIM-B (Table 15) was .21. Both were lower than for correlations with Meeting #1. This drop invalidates the notion that Meeting #1 behavior was so heavily loaded with introductions that real-life interaction could not emerge until later. As groups met longer, they showed no greater tendency to behave in accordance with their preferences as expressed in pretest HIM-B profiles.

Table 5

Correlation of weighted pre-test

HIM-B with HIM-SS raw profiles

Group	Meeting #1		Meeting #2	
	Median	Range	Median	Range
1	.46	.18 to .71	.29	.22 to .34
2	.18	-.24 to .47	.18 (N=4)	.06 to .27
3	.23	-.20 to .72	.27 (N=4)	-.14 to .55
4	.29	-.07 to .81	.44	.19 to .71
5	.41	-.04 to .61	.17	.03 to .38
6	.54	.25 to .61	.16	.04 to .42
Overall	.42 (N=30)	-.24 to .81	.21 (N=28)	-.14 to .71
Homogeneous Groups	.30 (N=20)	-.24 to .81	.20 (N=18)	-.14 to .71
Heterogeneous Groups	.44 (N=10)	-.04 to .71	.25 (N=10)	.03 to .38

Situational Test Prediction of Within-Group Behavior

Pretest HIM-V4 cell scores were correlated with HIM-SS group behavior cell scores for each S for Meeting #1 and again for Meeting #2. Table 6 shows these correlations for raw profiles, and Table 7 shows them for standardized profiles. In both cases, the median correlation with Meeting #1 profiles was about .70. The HIM-V4 as a situational test seemed to account for a much larger share of the variance in actual group behavior than did either the HIM-B or the weighted HIM-B as self-report inventories. About 50% of the variance of observed HIM-SS behavior arose from the same sources as did rated HIM-V4 behavior. This figure was a substantial portion of the maximum possible proportion, 65%, mentioned earlier as the stability of rated HIM-SS behavior itself. The hypothesis of Goldstein, et al. (1966, p. 329) appeared to be confirmed in this study: prediction of within-group behavior was more accurate with a behavioral measure than with a psychometric device.

Tables 6 and 7 show that prediction of Meeting #1 behavior by the pretest HIM-V4 was no better for standardized profiles than for raw profiles. The explanation for this lack of improvement in correlation may be that the distribution of stimulus items in the HIM-V4 was a good enough approximation to the distribution of stimulus HIM cells in the raw HIM-SS profiles for Meeting #1. For Meeting #2, the correlation for standardized profiles was higher than for raw profiles. However, the overall median correlations for both the standardized and raw profiles were lower for Meeting #2 than for Meeting #1. The implication is that the prevailing style of talk was similar in the HIM-V4 situation and in Meeting #1, but that it changed in Meeting #2.

Standardized profiles made the stimulus configuration similar to the two situations being correlated. Using standardized profiles, correlations of HIM-V4 profiles with Meeting #2 and with Meeting #1 should be about the same, unless time in the groups produced changes in Ss response styles. Table 7 shows that little if any such therapeutic changes were observed overall. The median PPM correlation from Meeting #1 to Meeting #2 dropped from .70 to .61. An interesting difference was noted, however, between homogeneous and heterogeneous groups. The median correlation in homogeneous groups went down, from .66 to .48; in heterogeneous groups it went up from .70 to .77.

The implication of this divergence of correlations seems to be that more therapy was done in the homogeneous groups than in the heterogeneous groups. Members of the latter tended to respond to a set of stimuli about the same way late in the groups as they did early in the groups. Members of homogeneous groups were less likely to respond the same way late as they did early. Tentatively it seems that more therapeutic movement can be accomplished in the early stages of homogeneous groups than of heterogeneous groups. Some clinical impressions on this issue are offered later in this chapter.

Significance tests of differences between self-report and situational measures as predictors. Correlated means t-tests were run on the distributions of squared correlations for each of three pretest predictors with Meeting #1

Table 6

Correlation of pre-test HIM-V4

with HIM-SS raw profile

Group	Meeting #1 PPM Correlation Median	Range	Meeting #2 PPM Correlation Median	Range
1	.77 (N=4)	.37 to .84	.63 (N=4)	.48 to .80
2	.45 (N=3)	.31 to .54	.16 (N=3)	-.23 to .57
3	.67 (N=4)	.55 to .80	.39 (N=3)	.36 to .40
4	.79	.53 to .89	.25	-.06 to .39
5	.77	.19 to .79	.61	.42 to .86
6	.82	.41 to .89	.46	-.20 to .84
Overall	.72 (N=26)	.19 to .89	.42 (N=25)	-.23 to .86
Homogeneous (2, 3, 4 &6)	.70 (N=17)	.32 to .89	.36 (N=16)	-.23 to .84
Heterogeneous (1 & 5)	.77 (N=9)	.19 to .84	.61 (N=9)	.42 to .86

NOTE: Unless otherwise specified, each median is based on N=5 group members. Where N<5, it is because the HIM-V4 recording was inaudible, or a member was absent from the group meeting.

Table 7

Correlation of pre-test HIM-V4

with HIM-SS standardized profile

Group	Meeting #1 PPM Correlation		Meeting #2 PPM Correlation	
	Median	Range	Median	Range
1	.79 (N=4)	.32 to .93	.71 (N=4)	.57 to .89
2	.60 (N=3)	.54 to .65	.23 (N=3)	-.16 to .66
3	.68 (N=4)	.65 to .83	.67 (N=3)	.49 to .71
4	.65	.50 to .87	.42	-.13 to .53
5	.69	.32 to .78	.77	.55 to .93
6	.75	.39 to .81	.61 (N=3)	.48 to .69
Overall	.70 (N=26)	.32 to .93	.61 (N=23)	-.16 to .93
Homogeneous Groups	.66 (N=17)	.39 to .87	.48 (N=14)	-.16 to .71
Heterogeneous Groups	.70 (N=9)	.32 to .93	.77 (N=9)	.55 to .93

HIM-SS raw profiles. The three predictors were the HIM-B, the weighted HIM-B, and the HIM-V4.

All mean differences tested were significant beyond the P = .05 level. The weighted HIM-B was a significantly better predictor of HIM-SS behavior than was the HIM-B. The HIM-V4 was a significantly better predictor than either of the other two. Table 8 shows these results.

Postdiction

Self-report. Posttest HIM-B profiles were correlated with HIM-SS raw profiles for Meeting #1 and Meeting #2. Experience in a group might allow persons to better visualize their characteristic behavior in groups. If so, postdiction would yield higher correlations with actual behavior than would prediction. As Table 9 shows, this expectation did not hold true. Where the correlation would be expected to be most high, for posttest HIM-B with Meeting #2, it was actually lower than for Meeting #1. The overall median correlations were not appreciably different from the correlations of pretest HIM-B with HIM-SS as shown in Table 4. In both cases, differences between homogeneous and heterogeneous groups were not large.

Other data also indicate that it is not likely that group members, due to the group experience, were better able to visualize their characteristic behavior in groups. The test-posttest HIM-B correlations were no different for experimental subjects than for control subjects (see page 27 of this paper).

Situational test postdiction. The correlations of posttest HIM-V4 profiles with HIM-SS Meeting #1 and Meeting #2 raw profiles were prepared for one homogeneous group and one heterogeneous group. These correlations are presented in Table 10. Overall, the median correlation for both meetings was about the same, .73 and .70, for the 10 Ss. As noted earlier with regard to raw pretest HIM-V4 correlations (Table 6), the median correlation seemed to decrease much more for members of homogeneous groups than of heterogeneous groups. In postdiction, the median correlation for the five members of the homogeneous group dropped from .81 for Meeting #1, to .45 for Meeting #2; for the heterogeneous group it held constant at .71.

Comments concerning pretest HIM-V4 correlations apply also to posttest. The way individuals talked seemed to change more from Meeting #1 to Meeting #2 in homogeneous groups than in heterogeneous groups. There was no significant tendency for postdiction to be more accurate than prediction. There may have been a slight tendency, however, in Group 4. Predictive correlations of HIM-V4 with Meeting #1 and Meeting #2 behavior were .79 and .25 (Table 6); comparable correlations in postdiction were .81 and .45.

Clinical Observations

Composition effects. As mentioned in the design of this study, Chapter II, the six groups were composed in different ways without the experimenter being

Table 8

	Weighted HIM-B compared to HIM-B	HIM-V4 compared to HIM-B	HIM-V4 compared to Weighted HIM-B
Mean Differences	.108	.355	.257
t	2.680	6.703	5.843
D.F.	29	24	24
P	.012	<.001	.001

Table 9

Correlation of posttest HIM-B with

HIM-SS raw profiles

Group	Meeting #1 PPM Correlation		Meeting #2 PPM Correlation	
	Median	Range	Median	Range
1	.28	-.01 to .66	-.11	-.38 to .23
2	.06	-.16 to .33	-.18 (N=4)	-.68 to -.14
3	.11	-.67 to .54	.02 (N=4)	-.71 to .23
4	.13	-.37 to .29	-.01	-.31 to .33
5	.03	-.35 to .47	-.28	-.40 to .31
6	.25	-.37 to .53	-.23	-.62 to .24
Overall	.14 (N=30)	-.67 to .66	-.14 (N=28)	-.71 to .33
Homogeneous	.12 (N=20)	-.67 to .54	-.12 (N=19)	-.71 to .33
Heterogeneous	.18 (N=10)	-.35 to .66	-.20 (N=9)	-.40 to .31

Table 10

Correlations of posttest HIM-V4

with HIM-SS raw profiles

Group	Meeting #1 PPM Correlation		Meeting #2 PPM Correlation	
	Median	Range	Median	Range
4 (Homogeneous)	.81	.67 to .86	.45	.20 to .55
5 (Heterogeneous)	.71	.35 to .85	.71	.70 to .82
Overall	.73 (N=10)	.35 to .86	.70 (N=10)	.20 to .82

told which group was which. While observing Meeting #1 of each group, he guessed the composition, homogeneous or heterogeneous, within the first 15 to 20 minutes. The guesses regarding the particular type of homogeneity were correct for groups 1, 2, 3, and 5. Group 4 and Group 6 were reversed. That is, the observer guessed that Group 4 was high on Member Ratio, and that Group 6 was high on Work Ratio, when, in fact, the reverse was true.

The differences between heterogeneous groups and homogeneous groups were striking. Some of these differences were as follows:

1. The heterogeneous groups were nervous, as typified by low toleration for silence. After a five-second silence, they burst into nervous laughter. The homogeneous groups tolerated silence with comfort. They were also smooth, affable, and cooperative from the very beginning. The members joked easily with each other after only 5 or 10 minutes in Group 2, which was homogeneous on Quadrant 3, a measure of tendency toward friendly pairing.

2. The heterogeneous groups were characterized by talk in staccato, overlapping bursts during which two or more members tried to talk at once. The flow of talk in the homogeneous groups was much more even.

3. Each of the two heterogeneous groups spent considerable time focusing on the therapist and his role. In Group 1, the focus on him was one of curiosity, and probing for indication of what should happen in the group. In Group 5, the focus on the therapist was one of anger and frustration with his failure to make the group meeting worthwhile. Both groups were resistant to the leader's probes to urge them into Quadrant 4 interaction. By contrast, the homogeneous groups readily followed the leader's urging. He seemed to be accepted in the homogeneous groups as a member with a particular function. His role was not a substantial topic for discussion.

4. The group leader smiled more frequently in the homogeneous groups, leaned back in his chair in a casual manner, and verbally expressed more feelings of warmth in the homogeneous groups. In the heterogeneous groups, the leader was less casual. In Group 5 he was distinctly defensive about his role, and resorted to considerably more assertive (HIM level C) talk than in any other group. The leader's use of Confrontive (HIM level E) interaction was sharply lower for Group 5.

5. At the outset of heterogeneous groups, members did not receive support from each other as they introduced themselves. Each person seemed to pass about the group looking for some response from the others, and jumping at any support that was perceived. Early introductions in the homogeneous groups were less like monologues. They more often elicited back-and-forth comparisons of similarities among members.

Group 3 was not like the other homogeneous groups. Groups 2, 4, and 6 were homogeneously high on different indexes of cooperative interaction. Group 3 was homogeneously high on Assertive (HIM level C).

The members of Group 3 spontaneously talked in abrasive, challenging ways. They described themselves as having life styles of isolation. They expressed disaffection with situations around them and added mention of their coping by rejection or derogation of these situations. "Sour grapes" attitudes were frequently expressed.

In Group 6, which was composed homogeneously high on Member Ratio, Member 1 was deviant in that he was below average on Member Ratio. During the group meetings, Member 1 was focused on as topic person far more than was any other member of Group 6. In describing himself, he said that he was a "radical". The subject of the later focusing on him consisted of a barrage of efforts to either understand or change his description of himself.

Experience with Member 1 in Group 6 stimulates a hypothesis to be tested in future research. He may be called a "deficient deviant" member of the group in that he was low on an index of predicted behavior on which the other members were homogeneously high. He also scored high on HIM-B total score, which is generally a measure of talkativeness. So the hypothesis that emerges is, in an otherwise homogeneous group, a deficient deviant who is inclined to be talkative will become a topic person: (a) for a longer time than other members, (b) more often and for a longer time than will untalkative deviants, and (c) be the focal point of efforts to change the deficient deviant to conform to some characteristic the group members may ascribe to themselves.

Group 2 was described by the observer as a very smooth and worthwhile group overall. Member 5 in this group was deviant on his HIM-B profile in a direction toward which the leader urged the group. The group members were high on Quadrant 3, but Member 5 was average on Quadrant 3, and high in Quadrant 4, the direction in which the leader urged the group. He might be referred to as a "seed deviant". Some of the smoothness attributed to Group 2 may have been made possible by Member 5 functioning as a model of the behaviors the leader suggested. He was the first member to make a confrontive statement toward another. Future research on group composition effects should include some study of groups seeded with such deviants. This seed deviant did not become the focus of efforts to get him to change such as the deficient deviant in Group 6 did.

Prediction techniques. The HIM-B seemed by clinical observation to show considerably more validity than the statistical measures revealed.

The HIM-V4 provoked several different comments from the men tested. A few said they were able to imagine the group and the other members very vividly. Most, however, complained about the lack of visual cues about the other members. They said they were simply unable to really put themselves into the imaginary situation. There were also frequent comments about the unrealistic acting in the recorded vignettes.

Suggestions for Future Research

Improved self-report inventory. Improvements can probably be made on the HIM-B that could produce an instrument which may have predictive validity as

high as .50 or .60. The advantage of such a self-report, paper-and-pencil inventory would be its ease of administration and scoring compared to a situational test predictor such as the HIM-V4. Routine use of the HIM-V4 in a clinical setting would be difficult because a trained and certified HIM rater would have to rate each one.

The starting point for an improved HIM-B should be careful development of a pool of items which are behaviorally-worded descriptions of characteristic behaviors in each HIM cell. A larger, more diverse norm group than the 100 college students Hill (1965) used to develop the HIM-B should be administered this item pool. Item responses could then be factor- or clustered-analyzed to pinpoint the item that best characterizes each of the 16 cells. This analysis might even reveal meaningful clusters of cells that can be combined together to give a different number of scales rather than direct correspondence with the 16 HIM cells. The items could then be put together in a forced-choice format, eliminating the difficulty the present HIM-B has in discriminating different types of preferred behavior. Response sets such as social desirability, and yea-saying are very likely on the present HIM-B. A worthwhile precedent to follow in arranging the items and scaling is presented by Borgen, Weiss, Tinsley, Dawis, and Lofquist (1968, pp. 10-22), who developed scales for measuring Occupational Reinforcer Patterns.

A generalized form of this improved HIM-B should be developed which will have balanced representation of each HIM cell. The scoring should then be tailored to each particular situation to which prediction is to be made. Each person's cell scores should be weighted by the relative proportion of each cell expected to be used in the group which he is to enter. As shown in this study, such weighting seems to substantially increase predictive validity over that offered by a perfectly balanced instrument.

A moderator scale could be developed to increase the predictability of the HIM-B. Considerable range was observed for the predictive validity of both predictors across individuals. Development and use of such moderator scales is discussed by Ghiselli (1963), and Dunnette (1966, pp. 163-168). Possibly a short version of a moderator scale could be built into an improved HIM-B, and used to print out an indication of the confidence that could be placed in an individual's predicted score.

Improved situational tests. If an improved version of the HIM-V4 were ever to be developed, it would be very good to include visual cues. Distributing snapshots of each actor on the tape might be one effective way of providing visual cues. A picture could be taken of the entire group when it meets to act the script; video taping the vignettes might be even better. Each vignette should give more context, and more should be shown about each group member early in the HIM-V. Better vignettes should be developed, perhaps as excerpts from actual portions of groups. They should be acted by professional actors, or perhaps by a real on-going group. A testing period beyond about 45 minutes would probably become especially tedious for most subjects. The strain might be relieved, and the predictive validity increased by presenting more context, and by requiring fewer responses from the subject.

The situational tests should probably be administered to a subject alone, or with the understanding that he is being monitored. Testing en masse such as

done in this study caused subjects to appear self-conscious, and distracted by each other. An automated device should be developed which would turn on the recording for S's responses, and not record the stimulus items. This simple innovation would cut the rater's time roughly in half compared to what was required of the raters in this study, who had to hear both stimulus and response for each item on each person's HIM-V4. The distribution of stimulus items in the HIM-V4 was rather arbitrarily set. An improved version should have the distribution set closer to what is expected in real life in the group meetings.

A situational test could be developed to contain some very explicit stimuli and some very vague stimuli. One type might prove to be more predictive than the other. For example, the ambiguous stimuli might provoke S to abandon his dependency on the external situation and revert to his characteristic repertoire. In this case, the ambiguous stimuli might be the better predictors. On the other hand, unambiguous stimuli might be clearly better predictors for persons who tend to be more field-dependent, in that they never do respond out of context. Field dependent persons might tend to imagine vivid situations to ambiguous stimuli, but these fantasies may not correspond to their actual behavior in live groups.

Modification in HIM-SS ratings. The HIM-SS can probably be improved as a reliable rating instrument. Some of the supplementary conventions developed for this study might be helpful. In particular, nothing to date has been published concerning the mechanics of how to rate, i.e., how large a sample of talk to consider in giving a rating, how many ratings to give within a long monologue, and whether to separate or ignore simultaneous statements by two or more group members. Additional sources of variance in individual behavior could probably be tapped by some expansion of the HIM framework. In particular a conspicuous and ostensibly reliable difference between people lies in their taking of the initiator or responder role. For example, some persons repeatedly respond to statements like, "I have had a lot of trouble with personal problems lately", with statements about themselves, like, "Yeah, me too. Like I've been having these bad dreams, see, and . . . ". Other persons consistently respond with statements like, "Oh, really? Tell me more about what has been on your mind". The latter response is an example of the initiator role, asking more about the other person. The responder role consists of telling more about yourself, rather than asking about the other person.

Wolf (1968) reviewed the literature on sequential interpersonal behavior and concluded that message sending and message receiving were two pervasive dimensions. These dimensions seem analogous to what were termed the initiator and responder roles above. The distinction could be across the 16 cells, and perhaps add a significant new measure of individual behavior.

Another meaningful distinction that could be added to HIM-SS ratings might be directness vs. indirectness. Consider, for example, a probe in cell IVD: "Howard, could you share with Bob here the feelings and impressions that you have of him right now?" A direct response to this probe would be, "Well, Bob, I don't know you too well, but I think I feel fairly positive toward you and I could get to know you better." An example of an indirect response would be, "Well I think Bob is a good guy, I like him okay, and would like to get to know him better." The distinction is that in a direct statement, the speaker addresses the person about whom he is talking. In an indirect statement, he talks about the other person without addressing him directly.

Trainability. The focus in this study was on prediction of verbal behavior. A logical extension of the methodology here would allow prediction of those who benefit from the type of treatment employed in the interpersonal skills training groups. Instead of using observed behavior frequencies as criterion, gain scores on desirable behavior indexes (for example HIM-SS Work Ratio) could be correlated with predictive profiles.

Long-term homogeneous and heterogeneous groups. Gross (1959) suggested that homogeneous and heterogeneous groups be conducted for longer than one hour to see how their differences persist. The present study found differences persisting through the fifth hour of group interaction. Both the observer and the group leader felt that the homogeneous groups were more productive during those five hours. But both also felt that, near the end, the homogeneous groups may have been reaching a plateau, while the heterogeneous groups were just entering a phase of rich productivity in which the diversity of styles would be an asset.

Summary

Purpose

The purpose of this study was to explore techniques for prediction of individual verbal behavior in small counseling groups. Specifically, the study engaged the problem of comparing a situational test and a self-report inventory as predictors of individual behavior.

Literature Review

The highest correlations with group behavior of individuals were obtained from objective measures of individual behavior in earlier groups. Correlations of ratings on single indexes of group-to-group behavior for individuals ranged from .50 to .90 (Bell and French, 1955; Borgatta and Bales, 1955a).

Prediction of within-group behavior of individuals based on their measured personality traits produced moderate or zero correlations (Breer, 1960; Derr and Silver, 1962; Mann, 1959).

The composition of groups has also been said to substantially affect the behavior of individuals in the groups. Some workers (Gradolf, 1958; Gross, 1959) have reported that group members only expressed behavior of their preferred kind in groups that were homogeneously high on that type of behavior. The effects of group composition were said to be particularly strong in groups that were composed on variables highly related to the assigned task of the group.

In prediction using situational tests, higher validity is generally obtained when three principles are observed: (a) consistency, (b) relation to task, and (c) objective observation.

In a pilot study, a situational test and a self-report inventory were compared as predictors of individual verbal styles in small groups. Both yielded profiles that correlated only about .10 with profiles of rated behavior in a group. Need was shown for (a) a skilled leader to stimulate interaction, (b) a longer group life than one hour, and (c) presentation of predictor stimulus sets similar in number and in kind of stimulus sets expected in criterion groups.

Design and Methodology

The system selected for verbal interaction analysis in this study also established the framework within which predictive measures were to be conceptualized. The system selected was the Hill Interaction Matrix (HIM). The system used for categorizing the talk in a group, statement-by-statement is termed the HIM-SS. Hill (1965) has also developed a self-report inventory, the HIM-B, for predicting individual behavior preferences in a small group. For the purpose of this study, a situational test was developed to predict an individual's verbal behavior in the 16 cells of the HIM. This situational test was the HIM-V4, a tape-recorded simulation of a small group meeting in which each subject was instructed to imagine himself as a member participating at designated intervals. What the individual said at each interval was recorded and later rated on the HIM-SS.

After 83 Ss were tested, 30 were selected to meet in criterion groups. There were six such groups, each consisting of five men, and each led by an experienced leader. Members of four of the six groups were homogeneously high on some index in their HIM-B profiles. Four different indexes were used: Quadrant 3, Row C (Assertive), Work Ratio, and Column IV (Relationship). The other two groups were formed of members who had no HIM-B index high in common. These latter two groups were termed heterogeneous. Him-B rather than HIM-V4 scores were used for composing the groups because the HIM-V4 could not be scored rapidly enough. HIM-B scores were also more nearly like the self-report measure which earlier studies (Gradolph, 1958; Gross, 1959) had used as composition variables. Different group composition effects may have been observed if the groups had been composed on situational test (HIM-V4) indexes rather than on self-report (HIM-B) indexes.

Each group met for a total of five hours in one week, in two 2-1/2 hour sessions. Each group meeting was tape-recorded on one track of a four-track stereo tape recorder. The observer attempted to make clinical judgements of the composition of each group. After all group meetings were completed, all control and experimental Ss were retested on the HIM-B and HIM-V4. The reason for the retesting was to get a measure of test-retest stability of these instruments, and also to provide profiles for postdiction to group behavior.

Statistical analyses consisted of correlations of individual profiles in predictor and criterion situations. The correlations calculated were product-moment correlations of type O in Cattell's (1952) notation. Each correlation was based on 16 variables (the scores in the 16 cells of the HIM) measured on two occasions, for one person.

Both raw score and standardized profiles were correlated. The standardization procedure was one developed specifically for this study, to get a set of stimuli which was equivalent in frequency and kind, in the predictor and criterion situations for each S.

Results and Conclusions

HIM-SS criterion behavior profiles for Meeting #1 were correlated with Meeting #2 for each S, to get a measure of the stability of the criterion. Raw profiles had a median correlation of .50; for standardized profiles it was .82. Early group behavior was presumed to be the best predictor of later group behavior, and the two situations had about 65% of their variance in common. Consequently, no predictor could be expected to correlate more highly than .82 with HIM-SS behavior.

The situational test showed significantly higher validity for predicting profiles of criterion behavior than did the self-report test in this study. The median correlation of the HIM-V4 (the situational test) with criterion behavior across 26 Ss tested on the pretest and participating in Meeting #1, was .72. The corresponding median correlation for the HIM-B (the self-report device was .03.

Each HIM-B cell score was weighted according to the relative frequency of presentation of that HIM cell as a stimulus in the HIM-V4. When these weighted HIM-B profiles were correlated with HIM-SS Meeting #1 profiles, the median was .42. Apparently part of the reason for the higher correlation of HIM-V4 profiles with HIM-SS profiles was that HIM-V4 scores had greater variance than did HIM-B scores. This finding implies that self-report predictive instruments can show higher validity coefficients for predicting the group behavior pattern of an individual if scores in the predictor categories are weighted according to their expected frequencies in the group.

The raw profiles of both predictors had lower correlations with Meeting #2 than with Meeting #1. Therefore, within-group behavior did not evolve into greater congruence with tested behavior. However, HIM-V4 correlations with standardized profiles were nearly the same for Meetings #1 and #2. These correlations led to the conclusion that therapeutic effects in the groups could not have accounted for more than about 17% of the variance in late-group behavior of individuals. The greater discrepancy for raw profiles of Meetings #1 and #2 showed that a prevailing style of talk in a group can obscure verbal behavior arising from individual preferences. Standardizing procedures are needed to correct for this obscuring. However, talk patterns may be largely caused by forces other than preferences. In particular, a strong tendency was noted for group members to conform to the prevailing talk categories in the groups. The rather high correlations of HIM-V4 with HIM-SS profiles should be interpreted with caution until the extent of these social conformity pressures is determined. One way to determine the extent would be to look at the intercorrelations of group behavior profiles of all possible pairs of members in a group. A high median intercorrelation would imply that most members are responding to a common source of stimulation, which may be other than their individual propensities.

For neither predictor were correlations higher for homogeneous groups than for heterogeneous groups. The findings of Gradolph (1958) and Gross (1959) were not supported in this study. However, an active leader was used in the present study which was not the case in Gradolph's or Gross' study.

Postdiction to within-group behavior did not give correlations appreciably different from prediction, for either the HIM-B or the HIM-V4. Experience in a group did not seem to improve the congruence of tested and real-life behaviors.

HIM-SS ratings were made on both HIM-V4 and group meeting tapes by three trained and certified HIM raters. Inter-rater reliability was determined by running all possible pairs of correlations between raters on tapes which all three had rated. The median of these 42 pairs was .70; the range was .20 to .98. These correlations were slightly lower than those reported elsewhere for HIM raters (Gibson, 1970; Hill, 1965).

The HIM-V4 had a test-retest reliability, over four weeks, of .91. For the HIM-B it was .56. Weighting the HIM-B cell scores like the HIM-V4 boosted the stability to .88. This increase reflected the greater variance of scores possible on the weighted HIM-B.

Some striking clinical observations were made by the observer of the groups. He was able to correctly deduce the composition of most groups after only 15 to 20 minutes of observation of Meeting #1 of each group. The heterogeneous groups were distinctly different from the homogeneous groups, being characterized by nervousness, strain, frustration, and expressions of hostility toward the leader. The homogeneous groups, by contrast, were smooth, affable, and cooperative, and readily followed the leader's suggestions. The observer felt that the homogeneous groups were more productive in the early hours, but that perhaps if continued longer, the heterogeneous groups may have provided a richer source for interaction.

Some suggestions for future research are the following:

1. Some groups should be composed homogeneously and others heterogeneously on HIM-B or HIM-V4 indexes, and conducted for longer than five hours. Four-man groups would probably provide optimum density of talk per person.

2. A self-report, paper-and-pencil instrument superior to the HIM-B should be developed.

3. A scale in the form of a moderator variable should be developed to differentiate predictable subjects from unpredictable subjects.

4. Instead of predicting profiles of behavior in a group, the data from a study like this could be used to predict trainability, which could be measured as improvement in the use of certain preferred verbal styles as the group progresses.

5. The effects of social conformity pressures on verbal behavior in counseling groups should be studied.

6. Any further work done with a situational test as a predictor should increase realism by including visual cues and more context about the simulated group than was used in the HIM-V4.

References

Amidon, E.J. and Hough, J.B. (Eds.). Interaction analysis: Theory, research, and application. Reading, Mass.: Addison-Wesley, 1967.

Anastasi, A. Psychological testing (3rd. ed.), New York: Macmillan, 1968.

Anderson, A.R. Group counseling. Review of Educational Research, 1969 (April), 39, (2), pp. 209-226.

Argyris, C. Unintended consequences of rigorous research. Psychological Bulletin, 1968, (September), 70, (3), 185-197.

Bales, R.F. Interaction process analysis: a method for the study of small groups. Cambridge, Mass.: Addison-Wesley, 1950.

Bell, G.B. and French, R.L. Consistency of individual leadership position in small groups of varying membership. In A.P. Hare, E.F. Borgatta, and R.F. Bales (Eds.), Small groups, New York: Alfred A. Knopf, 1955, pp. 275-280.

Bennis, W., Burke, R., Cutter, H., Harrington, H. and Hoffman, J. A note on some problems of measurement and prediction in a training group. Group Psychotherapy, 1957, 10, 328-341.

Blake, R.R., Mouton, Jane S., and Fruchter, B. The consistency of interpersonal behavior judgments made on the basis of short-term interactions in three-man groups. Journal of Abnormal and Social Psychology, 1954, 49, 573-578.

Borgatta, E.F., and Bales, R.F. The consistency of subject behavior and the reliability of scoring in interaction process analysis. In A.P. Hare, E.F. Borgatta, and R.F. Bales (Eds.) Small groups, New York: Alfred A. Knopf, 1955a, pp. 300-304.

Borgatta, E.F. and Bales, R.F. Interaction of individuals in reconstituted groups. In A.P. Hare, E.F. Borgatta, and R.F. Bales (Eds.), Small groups, New York: Alfred A. Knopf, 1955b, pp. 379-396.

Borgen, F.H., Weiss, D.J., Tinsley, H.E.A., Dawis, R.V., and Lofquist, L.H. The measurement of occupational reinforcer patterns. Minnesota Studies in Vocational Rehabilitation, 1968, 25.

Breer, P.E. Predicting interpersonal behavior from personality and role. Unpublished doctoral dissertation, Harvard University, 1960. Cited by A.P. Goldstein, K. Heller, and L.B. Sechrest, Psychotherapy and the psychology of behavior change, New York: Wiley, 1960, p. 332.

Cattell, R.B. The three basic factor-analytic research designs -- their interrelations and derivatives. Psychological Bulletin, 1952, 49, 499-520.

Cecil, E.A. The effects of group composition on the level of risk in group decisions. (Doctoral dissertation, University of Indiana) Ann Arbor, Mich.: University Microfilms, 1968. No. 68:4470. (Dissertation Abstracts, 1968, 28, (10), 3821A).

Derr, J. and Silver, A.W. Predicting participation and behavior in group therapy from test protocols. Journal of Clinical Psychology, 1962, 18, 322-325.

Dunnette, M.D. Personnel selection and placement. Belmont, California: Wadsworth Publishing Co., 1966.

Ghiselli, E.E. Moderating effects and differential reliability and validity. Journal of Applied Psychology, 1963, 47, 81-86.

Gibson, D.L. The HIM-G as a rapid method for verbal interaction analysis in a small group. Paper presented at the meeting of the American Personnel and Guidance Association, New Orleans, March, 1970.

Goldstein, A.P., Heller, K., and Sechrest, L.B. Psychotherapy and the psychology of behavior change. New York: Wiley, 1966.

Gradolph, I. The task approach of groups of single-type and mixed-type valency compositions. In Dorothy Stock and H.A. Thelan. Emotional dynamics and group culture. New York: New York University Press, 1958.

Gross, R.L. Therapy group composition: personal-interpersonal variable. (Doctoral dissertation, University of Utah). Ann Arbor, Mich.: University Microfilms, 1959, No. 59-6075.

Helmstadter, G.C. Principles of psychological measurement. New York: Appleton-Century-Crofts, 1964.

Hill, W.F. Hill Interaction Matrix Scoring Manual. Los Angeles: Public Systems Research Institute, University of Southern California, 1963.

Hill, W.F. Hill Interaction Matrix: a method for studying interaction in psychotherapy groups. Los Angeles: Public Systems Research Institute, University of Southern California, 1965.

Hill, W.F. A supplement to the Hill Interaction Matrix monograph. Los Angeles: Public Systems Research Institute, University of Southern California, 1966.

Jackson, D.N. Personality Research Form Manual. Goshen, N.Y.: Research Psychologists Press, 1967.

Kelly, E.E. Group counseling interaction and member personality. (Doctoral dissertation, University of Texas.) Ann Arbor, Mich.: University Microfilms, 1965. No. 65-4286 (Dissertation Abstracts, 1965, 26, (9), 5117).

McGrath, J.E. and Altman, I. Small group research, New York: Holt, Rinehart and Winston, 1966.

McNemar, Q. Psychological statistics. (3rd ed.) New York: Wiley, 1962.

Magnusson, D., Gerzen, Marianne, and Nyman, B. The generality of behavioral data I: generalization from observations on one occasion. Multivariate Behavioral Research, 1968 (July), 3, (3), pp. 295-320.

Mann, R.D. The relation between personality characteristics and individual performance in small groups. (Doctoral dissertation, University of Michigan, 1959) Ann Arbor, Mich.: University Microfilms, 1959, No. 59-2153. (Dissertation Abstracts, 1959, 19, (12), 3393-3394).

Roffers, T. Conditioning a style of interaction in counseling groups and its effects on behavior and attitude changes in college students. Unpublished doctoral dissertation, University of Minnesota, 1969.

Schutz, W.C. FIRO: a three-dimensional theory of interpersonal behavior. New York: Holt, Rinehart and Winston, 1960.

Schutz, W.C. The interpersonal underworld (Reprint edition). Palo Alto, Calif.: Science and Behavior Books, 1966.

Seligman, M. and Sterne, D.M. Verbal behavior in therapist-led, leaderless, and alternating group psychotherapy sessions. Journal of Counseling Psychology, 1969, 16, 325-328.

Stager, D.P. Conceptual level as a composition variable in small group decision-making. (Doctoral dissertation, Princeton University, 1966). Ann Arbor, Mich.: University Microfilms, 1966. No. 66-13355. (Dissertation Abstracts, 1966, 27, (6), 1932-A).

Stock, Dorothy. A survey of research on T-groups. In L.P. Bradford, J.R. Gibb, and K.D. Benne (Eds.) T-group theory and laboratory method. New York: Wiley, 1964, pp. 395-441.

Thorn, J.P. Prediction of behavior in a group. Paper presented at the Meeting of the American Personnel and Guidance Association, New Orleans, March, 1970.

Toobert, S. The relation between personality and interaction behavior in small groups. (Doctoral dissertation, University of Oregon). Ann Arbor, Mich.: University Microfilms, 1966, No. 66-635. (Dissertation Abstracts, 1966, 26, (9), 5574-5575).

Weislogel, R.L. and Schwarz, P.A. Some practical and theoretical problems in situational testing, Educational and Psychological Measurement, 1955, 15, pp. 39-46

Wolf, G. Dynamics of cooperative and persuasive interpersonal behavior. Ann Arbor, Mich.: University Microfilms, 1968, No. 68-680. (Dissertation Abstracts, 1968, 28, (7), 2778A).

APPENDIX A

Computer-scored HIM-B profile

>>>>>>>> MILL INTERACTION MATRIX (MIM-B) SCORES AND NORMS <<<<<<<<

IDENTIFICATION NUMBER= 01503811750216

DATE SCORED WAS 02/27/70
ANSWER SHEET RESPONSE RANGE = 0 TO 5
GROUP NUMBER = 015
SEQUENTIAL SUBJECT NUMBER= 038
SEX= MALE
TESTING PERIOD= 1
STUDENT I.D. NUMBER= 750216

	TOPIC I	GROUP II	PERSONAL III	RELNSHIP IV	RAW SCORE	PERCENT	NORM
CONVENTIONAL-B	1230 =6	1204 =7	1234 =10	1034 =8	31	27.0	=
ASSERTIVE-C	1204 =7	1030 =4	1000 =1	1000 =1	13	11.3	0
SPECULATIVE-D	1034 =8	1234 =10	1234 =10	1204 =7	35	30.4	•
CONFRONTIVE-E	1234 =10	1230 =6	1234 =10	1234 =10	36	31.3	•
RAW SCORE	31	27	31	26			
PERCENT	27.0	23.5	27.0	22.6			
NORM	•	0	0	0			

TOTAL SCORE= 115
TOTAL SCORE NORM = •

CELL WEIGHTS

	I	II	III	IV
B	1	2	9	10
C	3	4	11	12
D	5	6	13	14
E	7	8	15	16

RISK RATIO =	.742	=
WORK RATIO =	1.614	•
INTRA-GROUP RATIO =	.855	0
MEMBER RATIO =	.983	•
MEAN ACCEPTED CELL =	8.748	0

0/0 OF TOTAL POSSIBLE WEIGHTED SCORES = 74.8 %

QUADRANT ANALYSIS

QUAD 1		QUAD 3	
RAW SCORE=	24	RAW SCORE=	28
PERCENT=	20.9	PERCENT=	17.4
NORM=	-	NORM=	0
RISK RATIO=	.846	RISK RATIO=	.111
QUAD 2		QUAD 4	
RAW SCORE=	34	RAW SCORE=	37
PERCENT=	29.6	PERCENT=	32.2
NORM=	•	NORM=	•
RISK RATIO=	.889	RISK RATIO=	1.176

APPENDIX B

Interpretation of the HIM-B profile

Interpreting Your HIM-B Profile

The Hill Interaction Matrix (HIM) categorizes a person's talk in two ways: first, what he talks about, and second, how he talks about it. When you filled out the HIM-B questionnaire, you were describing your own preferences for what you talk about in a group, and how you talk about it.

The "What" dimension is shown by roman numerals. The safest thing you can talk about in a group is I, a topic of general interest, like the weather, politics, psychology, etc. Next, you can talk about the Group itself (II). Next you can talk about yourself, or you can participate in conversation that focuses on another group member who is topic person. Such conversation is called Personal (III) because it focuses on one person present. The most risky thing to talk about, from the standpoint of your vulnerability to embarassment in the group, is IV, a Relationship in the here-and-now, between two or more persons in the group, one of whom could be you.

The way you prefer to talk, the "How" dimension, is shown by B,C,D, and E. Moving down on this dimension indicates an openness to changing your opinions, attitudes, and characteristic behavior. Changes like that require effort, so this "How" dimension is a "Work" scale. The least effortful way to talk is B, Conventional. This is routine socializing, small talk, and where-are-you-from information-seeking. It takes only a little more effort to be Assertive (C). This is how you talk when you argue, gripe, blow off steam, tell someone off, or try to persuade. At neither B or C are you willing to change anything about yourself. You begin to be open to change when you begin to think about it. This thoughtfulness is reflected in D, the Speculative way of talking. The Confrontive style, E, is the hardest work. It involves honesty, insight, taking responsibility for what you say by using specific examples, and getting down to the real core of the issue at hand.

On the HIM-B printout, the "What" and "How" dimensions intersect to form 16 cells. The highest score you can get in any cell is 10, and the highest sum for any column or row is 40. Each column and row score is converted to a percentage of the total raw score. Your percentages were compared against hundreds of other persons like you to see how you stack up. In the places labeled "NORM", a zero means you prefer to talk about that subject (for the norms at the end of each row) about as much as the average person. A minus means you prefer that style less than do most persons like you. A plus sign indicates your preference is greater than average.

The Total Score, shown in the box, is a measure of your general talkativeness. The Total Score Norm right under it shows whether you think you talk more, less, or about the same as others in a group.

The Risk Ratio is the sum of rows C and E, divided by the sum of B and D. That is, RR = (C+E)/(B+D). The higher this number is, the more likely you are to go out on a limb in the way you talk, rather than playing it safe. You say that you readily risk being put down, contradicted, embarrassed, or rejected in a group, if you got a plus sign after Risk Ratio.

Your Work Ratio is calculated by (D+E)/B+C). It measures your openness to being influenced by others, and to exerting helpful influence on them. A plus sign here means that you like to get down to business more than the average group member. A minus sign indicates you tend to avoid talking in a way that may require you to change the way you are.

Your Intra-Group Ratio is calculated by (II + IV)/(I = III). Its meaning is rather obscure and probably not too helpful to you. It represents your tendency to talk about the group and about relationships among members, rather than talking about general interest topics or individual group members. A person sophisticated in group dynamics would be likely to get a plus sign here. He probably would function better in a sensitivity training T-group conducted for the benefit of an organization, than he would in a therapy, counseling, or skills group conducted primarily for the members to increase their self-understanding or to solve personal problems.

Member Ratio is (III + IV)/(I + II). A high number here would indicate your preference to explore the people present in the group, and their immediate relationships with each other, rather than talk about things that are less personal. A minus sign after Member Ratio indicates that you tend to avoid getting close to others, or letting them know you intimately.

The remaining material in the printout is for research use only. The everyday significance of these measures is not yet known.

By way of conclusion, look back at your matrix of cell scores. A normal, well-functioning person should have a pretty well balanced profile. He should have few if any zero cell scores, and certainly not all tens. Such a person is capable of interacting with others in a variety of ways, and is probably able to judge when each style is appropriate. There is a definite place for each style in ordinary human relationships. If there are one or more columns or rows where you had a couple of zero cell scores, and a minus sign for your norm, you might want to seek out some sort of group experiences designed to strengthen that underdeveloped aspect of your interpersonal behavior.

APPENDIX C

Printouts of standardized

HIM-V4 and HIM-SS

profiles for one person

DATE RUN= 06/08/70 SIU MIN-VA PROFILE FOR I.D.= 750216 MEMBER NUMBER= 1 IN GROUP= 1

RATER NUMBER= 2 RATING OCCASION = 2 TESTING OCCASION= 1

MIN CELL OF RESPONSE # \ MIN CELL OF STIMULUS #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	RESPONSE TOTALS
1			1						1								2
2																	
3																	0
4																	
5					1	1		1	1				1	2	1	1	9
6						1		1									2
7																	
8																	0
9								1	4				1	3		1	10
10										2			1	5			8
11																	
12																	
13													3				3
14										1				7			8
15																	0
16																	0
17																	
STIMULUS TOTALS #			1		1	2		3	6	3			6	17	1	2	42

RISK RATIO= .214 WORK RATIO= .759 MEMBER RATIO= 9.231 INTRA-GROUP RATIO= .750

MEAN RESPONSE CELL# 9.8688 SUMMED WEIGHTED RESPONSES= 380 TOTAL NUMBER OF RESPONSES= 42

DIAGONAL INDEXES: FREQUENCY OVR 18 FREQUENCY ABOVE (CLOSED)= 22 FREQUENCY BELOW (OPEN)= 2

ABOVE-DIAGONAL (CLOSED) DISCREPANCIES: SUM= 116 AVERAGE= 5.273

BELOW-DIAGONAL (OPEN) DISCREPANCIES: SUM= 5 AVERAGE= 2.500

RATIOS: OPENNESS RATIO (ORAT)= .043 OPENNESS DISCREPANCY RATIO (ODRI)= .474 OPENNESS FREQUENCY RATIO (FRAT)= .091

DIFFERENCES: NET FREQUENCY BELOW (NFT)= -20 NET DISCREPANCY BELOW (ONFT)=-111

DATE RUN= 06/04/70

STIMULUS-SET PROFILE FOR MEMBER 1, ID= 750216 GROUP NUMBER 1 MEETING NUMBER 1
TIME PERIOD= 0 TO 0 MINUTES

RATER NUMBER= 2 RATING OCCASION = 1

MIM CELL OF RESPONSE =	MIM CELL OF STIMULUS = 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	RESPONSE TOTALS
1														1			1
2																	0
3														1			3
4					1					1						1	4
5						1		2									1
6							1							2			8
7									6					6			6
8										2				1			2
9														1			1
10														1			8
11	1												6	1	1		9
12														6		1	8
13									1								1
14										•							0
15														2			2
16																	0
17																	0
STIMULUS TOTALS =	1	1	2	3	6	3	6	17	1	2							42

RISK RATIO= .250 WORK RATIO= 1.143 MEMBER RATIO= 3.667 INTRA-GROUP RATIO= 1.000

MEAN RESPONSE CELL= 10.457 SUMMED WEIGHTED RESPONSES= 435 TOTAL NUMBER OF RESPONSES= 42

DIAGONAL INDEXES: FREQUENCY ON= 20 FREQUENCY ABOVE (CLOSED)= 17 FREQUENCY BELOW (OPEN)= 5

ABOVE-DIAGONAL (CLOSED) DISCREPANCIES: SUM= 76 AVERAGE= 4.471

BELOW-DIAGONAL (OPEN) DISCREPANCIES: SUM= 20 AVERAGE= 4.000

RATIOS: OPENNESS RATIO (ORAT)= .263 OPENNESS DISCREPANCY RATIO (ODR)= .895 OPENNESS FREQUENCY RATIO (FRAT)= .294

DIFFERENCES: NET FREQUENCY BELOW (NET)= -12 NET DISCREPANCY BELOW (DNET)= -56

OFFICE OF NAVAL RESEARCH

PERSONNEL AND TRAINING RESEARCH PROGRAMS (Code 458)

DISTRIBUTION LIST

Contract No. N00014-68-A-0141-0003

Contractor University of Minn. - Dunnette

Navy

4 Chief of Naval Research
Code 458
Department of the Navy
Washington, D.C. 20360

1 Director
ONR Branch Office
495 Summer Street
Boston, Massachusetts 02210

1 Director
ONR Branch Office
219 South Dearborn Street
Chicago, Illinois 60604

1 Director
ONR Branch Office
1030 East Green Street
Pasadena, California 91101

6 Director, Naval Research Laboratory
Washington, D.C. 20390
ATTN: Library, Code 2029

1 Office of Naval Research
Area Office
207 West 24th Street
New York, New York 10011

1 Office of Naval Research
Area Office
1076 Mission Street
San Francisco, California 94103

6 Director
Naval Research Laboratory
Washington, D.C. 20390
ATTN: Technical Information Division

12 Defense Documentation Center
Cameron Station, Building 5
5010 Duke Street
Alexandria, Virginia 22314

1 Commanding Officer
Service School Command
U.S. Naval Training Center
San Diego, California 92133

3 Commanding Officer
Naval Personnel and Training
Research Laboratory
San Diego, California 92152

1 Commanding Officer
Naval Medical Neuropsychiatric
Research Unit
San Diego, California 92152

1 Commanding Officer
Naval Air Technical Training Center
Jacksonville, Florida 32213

1 Dr. James J. Regan, Code 55
Naval Training Device Center
Orlando, Florida 32813

1 Dr. Martin A. Bertin
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 08540

1 Technical Library
U.S. Naval Weapons Laboratory
Dahlgren, Virginia 22248

1 Research Director, Code 06
Research and Evaluation Department
U.S. Naval Examining Center
Building 2711 - Green Bay Area
Great Lakes, Illinois 60088

1 Dr. A.L. Slafkosky
Scientific Advisor (Code AX)
Commandant of the Marine Corps
Washington, D.C. 20380

1 Behavioral Sciences Department
Naval Medical Research Institute
National Naval Medical Center
Bethesda, Maryland 20014

1 Commanding Officer
Naval Medical Field Research Laboratory
Camp Lejeune, North Carolina 28542

1 Deputy Director
Office of Civilian Manpower Management
Department of the Navy
Washington, D.C. 20390

1 Director
Aerospace Crew Equipment Department
Naval Air Development Center
Johnsville
Warminster, Pennsylvania 18974

1 Chief
Naval Air Technical Training
Naval Air Station
Memphis, Tennessee 38115

1 Director
Education and Training Sciences Dept.
Naval Medical Research Institute
National Naval Medical Center
Building 142
Bethesda, Maryland 20014

1 Commander
Submarine Development Group TWO
Fleet Post Office
New York, New York 09501

1 Commander
Operational Test & Evaluation Force
U.S. Naval Base
Norfolk, Virginia 23511

1 Mr. S. Friedman
Special Assistant for Research & Studies
OASN (M&RA)
The Pentagon, Room 4E794
Washington, D.C. 20350

1 Chief of Naval Operations, (Op-07TL)
Department of the Navy
Washington, D.C. 20350

1 Chief of Naval Material (MAT 031M)
Room 1323, Main Navy Building
Washington, D.C. 20360

1 Mr. George N. Graine
Naval Ship Systems Command (SHIP 03H)
Department of the Navy
Washington, D.C. 20360

1 Chief
Bureau of Medicine and Surgery
Code 513
Washington, D.C. 20390

1 Chief
Bureau of Medicine and Surgery
Research Division (Code 713)
Department of the Navy
Washington, D.C. 20390

6 Technical Library (Pers-11b)
Bureau of Naval Personnel
Department of the Navy
Washington, D.C. 20370

3 Personnel Research and Development Laboratory
Washington Navy Yard, Building 200
Washington, D.C. 20390
ATTN: Library, Room 3307

1 Commandant of the Marine Corps
Headquarters, U.S. Marine Corps
Code AO1B
Washington, D.C. 20380

1 Technical Library
Naval Ship Systems Command
Main Navy Building, Room 1532
Washington, D.C. 20360

1 Technical Library Branch
Naval Ordnance Station
Indian Head, Maryland 93940

1 Library, Code 0212
Naval Postgraduate School
Monterey, California 93940

1 Technical Reference Library
Naval Medical Research Institute
National Naval Medical Center
Bethesda, Maryland 20014

1 Scientific Advisory Team (Code 71)
Staff, COMASWFORLANT
Norfolk, Virginia 23511

3 Technical Director
Personnel Research Division
Bureau of Naval Personnel
Washington, D.C. 20370

1 Deputy
Office of Civilian Manpower
Management
Department of the Navy
Washington, D.C. 20390

1 Technical Library
Naval Training Device Center
Orlando, Florida 32813

1 Dr. Earl I. Jones
Director
Naval Training Research Institute
Naval Personnel & Training
Research Laboratory
San Diego, California

1 Head, Personnel Measurement Staff
Capital Area Personnel Service Office-
Navy
Ballston Tower #2, Room 1204
801 N. Randolph St.
Arlington, Virginia 22203

Army

1 Director of Research
U.S. Army Armor Human Research Unit
Fort Knox, Kentucky 40121
ATTN: ATSAG-EA

1 Director
Behavioral Sciences Laboratory
U.S. Army Research Institute of
Environmental Medicine
Natick, Massachusetts 01760

1 U.S. Army Behavior and Systems
Research Laboratory
Commonwealth Building, Room 239
1320 Wilson Boulevard
Arlington, Virginia 22209

1 Division of Neuropsychiatry
Walter Reed Army Institute of Research
Walter Reed Army Medical Center
Washington, D.C. 20012

1 Behavioral Sciences Division
Office of Chief of Research and
Development
Department of the Army
Washington, D.C. 20310

1 Commandant
U.S. Army Adjutant General School
Fort Benjamin Harrison, Indiana 46216
ATTN: ATSAG-EA

1 Dr. George S. Harker, Director
Experimental Psychology Division
U.S. Army Medical Research Laboratory
Fort Knox, Kentucky 40121

1 LTC William C. Cosgrove
USA CDC Personnel & Administrative
Services Agency
Ft. Benjamin Harrison, Indiana 46216

Air Force

1 Commandant
U.S. Air Force School of Aerospace
Medicine
ATTN: Aeromedical Library (SMSL-4)
Brooks Air Force Base, Texas 78235

1 AFHRL (TR/Dr. G.A. Eckstrand)
Wright-Patterson Air Force Base
Ohio 45433

1 Personnel Research Division (AFHRL)
Lackland Air Force Base
San Antonio, Texas 78236

1 AFOSR (SRLB)
1400 Wilson Boulevard
Arlington, Virginia 22209

1 Headquarters, U.S. Air Force
AFPTRBD
Programs Resources and Technology Div.
Washington, D.C. 20330

1 AFHRL (HRTT/Dr. Ross L. Morgan)
Wright-Patterson Air Force Base
Ohio 45433

1 Lt. Col. John E. Dulfer
HQ, AFSC (SDEC)
Andrews Air Force Base
Washington, D.C. 20330

DOD

1 LTCOL F.R. Ratliff
Office of the Assistant Secretary
of Defense (M&RU)
The Pentagon, Room 3D960
Washington, D.C. 20301

1 Dr. Ralph R. Canter
Military Manpower Research Coordinator
OASD (M&RA) MR&U
The Pentagon, Room 3D960
Washington, D.C. 20301

Government

1 Dr. Andrew R. Molnar
Computer Innovation in Education
Section
Office of Computing Activities
National Science Foundation
Washington, D.C. 20550

1 Dr. Alvin E. Goins, Executive Secretary
Personality and Cognition Research
Review Committee
Behavioral Sciences Research Branch
National Institute of Mental Health
5454 Wisconsin Avenue, Room 10A02
Chevy Chase, Maryland 20015

1 Director, National Center for
Educational Research & Development
U.S. Office of Education
Dept. of Health, Education & Welfare
Washington, D.C. 20202

1 Mr. Joseph J. Cowan, Chief
Psychological Research Branch (P-1)
U.S. Coast Guard Headquarters
400 Seventh Street, S.W.
Washington, D.C. 20226

Non-Government

1 ERIC Clearinghouse on Vocational
and Technical Education
The Ohio State University
1900 Kenny Road
Columbus, Ohio 43210
ATTN: Acquisition Specialist

1 ERIC Clearinghouse on
Educational Media and Technology
Stanford University
Stanford, California 94305

1 Dr. Don H. Coombs, Co-Director
ERIC Clearinghouse
Stanford University
Palo Alto, California 94305

1 Dr. Richard C. Atkinson
Department of Psychology
Stanford University
Stanford, California 94305

1 Dr. Richard S. Hatch
Decision Systems Associates, Inc.
11428 Rockville Pike
Rockville, Maryland 20852

1 Director
Human Resources Research Organization
300 North Washington Street
Alexandria, Virginia 22314

1 Human Resources Research Organization
Division #3, Recruit Training
Post Office Box 5787
Presidio of Monterey, California 93940

1 Human Resources Research Organization
Division #5, Air Defense
Post Office Box 6021
Fort Bliss, Texas 79916

1 Human Resources Research Organization
Division #4, Infantry
Post Office Box 2086
Fort Benning, Georgia 31905

1 Human Resources Research Organization
Division #6, Aviation
Post Office Box 428
Fort Rucker, Alabama 36360

1 Dr. Robert J. Seidel
Human Resources Research Organization
300 N. Washington St.
Alexandria, Virginia 22314

1 Dr. Marvin D. Dunnette
University of Minnesota
Department of Psychology
Elliott Hall
Minneapolis, Minnesota 55455

1 S. Fisher, Research Associate
Computer Facility
Graduate Center
City University of New York
33 West 42nd Street
New York, New York 10036

1 Dr. John C. Flanagan
American Institutes for Research
Post Office Box 1113
Palo Alto, California 94302

1 Dr. Robert Glaser
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, Pennsylvania 15213

1 Dr. Albert S. Glickman
American Institutes for Research
8555 Sixteenth Street
Silver Spring, Maryland 20910

1 Dr. Bert Green
Department of Psychology
Johns Hopkins University
Baltimore, Maryland 21218

1 Dr. Duncan N. Hansen
Center for Computer Assisted Instruction
Florida State University
Tallahassee, Florida 32306

1 Dr. M.D. Havron
Human Sciences Research, Inc.
Westgate Industrial Park
7710 Old Springhouse Road
McLean, Virginia 22101

1 Dr. Carl E. Helm
Department of Educational Psychology
Graduate Center
City University of New York
33 West 42nd Street
New York, New York 10036

1 Mr. Harry H. Harman
Division of Computation Sciences
Educational Testing Service
Princeton, New Jersey 08540

1 Dr. Lee J. Cronbach
School of Education
Stanford University
Stanford, California 94305

1 Psychological Abstracts
American Psychological Association
1200 Seventeenth Street, N.W.
Washington, D.C. 20036

1 Dr. Bernard M. Bass
University of Rochester
Management Research Center
Rochester, New York 14627

1 Dr. Lee R. Beach
Department of Psychology
University of Washington
Seattle, Washington 98105

1 Dr. Roger A. Kaufman
Graduate School of Leadership & Human Behavior
U.S. International University
8655 E. Pomerada Road
San Diego, California 92124

1 Dr. George E. Rowland
Rowland and Company, Inc.
Post Office Box 61
Haddonfield, New Jersey 08033

1 Dr. Mats Bjorkman
University of Umea
Department of Psychology
Umea 6, SWEDEN

1 Mr. Roy Ference
Room 2311
U.S. Civil Service Commission
Washington, D.C. 20415

1 Dr. Frederic M. Lord
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 08540

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
Santa Barbara Research Park
6780 Cortona Drive
Goleta, California 93017

1 Dr. Stanley M. Nealey
Department of Psychology
Colorado State University
Fort Collins, Colorado 80521

1 Dr. Gabriel D. Ofiesh
Center for Educational Technology
Catholic University
4001 Harewood Road, N.E.
Washington, D.C. 20017

1 Mr. Luigi Petrullo
2431 North Edgewood Street
Arlington, Virginia 22207

1 Dr. Len Rosenbaum
Psychology Department
Montgomery College
Rockville, Maryland 20852

1 Dr. Arthur I. Siegel
Applied Psychological Services
Science Center
404 East Lancaster Avenue
Wayne, Pennsylvania 19087

1 Dr. Paul Slovic
Oregon Research Institute
Post Office Box 3196
Eugene, Oregon 97403

1 Dr. Diane M. Ramsey-Klee
R-K Research & System Design
3947 Ridgemont Drive
Malibu, California 90265

1 Dr. Ledyard R. Tucker
University of Illinois
Psychology Building
Urbana, Illinois 61820

1 Dr. John Annett
Department of Psychology
Hull University
Hull
Yorkshire, England

1 Dr. Lloyd G. Humphreys
Assistant Director for Education
National Science Foundation
Washington, D.C. 20550

1 Dr. Joseph W. Rigney
Behavioral Technology Laboratories
University of Southern California
University Park
Los Angeles, California 90007

1 Educational Testing Service
Division of Psychological Studies
Rosedale Road
Princeton, New Jersey 08540